

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

CARL MURCHISON

GORDON ALLPORT  
Harvard University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
Smithsonian Institution

HERBERT S. CONRAD  
U. S. Office of Education

ELMER K. CULLER  
University of Rochester

JOHN F. DASHIELL  
University of North  
Carolina

ROY M. DORCUS  
University of California  
at Los Angeles

HORACE B. ENGLISH  
Ohio State University

FRANKLIN FEARING  
University of California  
at Los Angeles

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
University of Virginia

CLARENCE H. GRAHAM  
Columbia University

ERNEST R. HILGARD  
Stanford University

HEINRICH KLÜVER  
University of Chicago

CARNEY LANDIS  
Columbia University

NORMAN L. MUNN  
Bowdoin University

JOHN PAUL NAFE  
Washington University

DONALD G. PATERSON  
University of Minnesota

LEE EDWARD TRAVIS  
University of Southern  
California

ROBERT C. TRYON  
University of California

MICHAEL J. ZIGLER  
Wellesley College

## VOLUME 54

1962

Copyright, 1962, by The Journal Press  
Published Quarterly by The Journal Press  
Provincetown, Massachusetts, U. S. A.

# AUTHOR INDEX

Allison, Roger B., Jr. ....	491	Hendsch, Gene .....	241
Anderson, Larry M. ....	65	Hood, William Robert .....	121
Andronico, Michael P. ....	209	Horn, Paul .....	139
Aronoff, Joel .....	39	Horne, Wiliam C. ....	251
Asch, Solomon E. ....	3	Jenkins, Thomas N. ....	417
Barber, Theodore Xenophon .....	363	King, H. E. ....	299
Bartley, S. Howard .....	245	Knehr, Charles A. ....	165, 443, 467
Bell, Deanne M. ....	113	Korner, Ija N. ....	491
Bennis, Warren G. ....	391	Langer, Philip .....	203
Blackman, Sheldon .....	197	Lawton, M. Powell .....	131
Boardman, William K. ....	33	Loeb, M. ....	101
Butters, Nelson .....	473	Mulholland, Thomas .....	69, 317
Cahn, Byrton .....	131	Oltman, Philip K. ....	485
Cassel, Russell N. ....	241	Pasamanick, Benjamin .....	453
Calverley, David Smith .....	363	Pastore, Nicholas .....	293
Chalmers, Douglas K. ....	251	Raygor, Alton W. ....	85
Cohen, Arthur M. ....	391	Rickels, K. ....	331, 345
Compton, Norma H. ....	191	Rosenbaum, Milton E. ....	251
Crofoot, Alan P. ....	85	Rosenberg, Milton J. ....	485
Day, David R. ....	259	Runnals, Sylvia .....	69, 317
Downing, R. W. ....	321, 345	Schneider, Carl W. ....	245
Ebenholtz, Sheldon M. ....	3	Shapiro, Stewart B. ....	285
England, Samuel J. M. ....	453	Sherif, Muzafer .....	121
Fletcher, Mildred E. ....	47	Spigel, Irwin .....	221
Forgays, Donald G. ....	209	Stogdill, Ralph M. ....	259
French, Thomas M. ....	271	Wells, Hal M. ....	113
Fromm, Erika .....	271	Webb, Warren W. ....	159
Furst, Edward J. ....	85	Webster, Frances R. ....	159
Goldman, Samuel .....	229	Wiener, Morton .....	473
Goldstone, Sanford .....	33, 159	Wolfe, John B. ....	139
Goode, Omar S. ....	259	Wright, David J. ....	33
Gorfein, David S. ....	65	Zajac, J. L. ....	153
Hawkes, Glenn R. ....	101, 309	Zwanziger, Max. D. ....	491

## ERRATUM

In Volume 54, Page 131, the word "minutes," which appears three times, should be "seconds."



## TABLE OF CONTENTS

The process of free recall: evidence for non-associative factors in acquisition and retention . . . . .	3
BY SOLOMON E. ASCH AND SHELDON M. EBENHOLTZ	
Time judgment and schizophrenia: step interval as a relevant contextual factor . . . . .	33
BY DAVID J. WRIGHT, SANFORD GOLDSTONE, AND WILLIAM K. BOARDMAN	
Freud's conception of the origin of curiosity . . . . .	39
BY JOEL ARONOFF	
Attitudes of a group of nursing students toward a clinical affiliation in a public hospital for the mentally ill . . . . .	47
BY MILDRED E. FLETCHER	
A note on the validity of the Bernberg Human Relations Inventory . . . .	65
BY DAVID S. GORFEIN AND LARRY M. ANDERSON	
A stimulus-brain feedback system for evaluation of alertness . . . . .	69
BY THOMAS MULHOLLAND AND SYLVIA RUNNALS	
Basic motivation and concept of nursing as chosen profession . . . . .	85
BY EDWARD J. FURST, ALTON W. RAYGOR, AND ALAN P. CROFOOT	
Detection of differences in duration of acoustic and electrical cutaneous stimuli in a vigilance task . . . . .	101
BY M. LOEB AND G. R. HAWKES	
Binocular perceptual discriminations of authority and peer group figures among over, under and equal achievers . . . . .	113
BY HAL M. WELLS AND DEANNE M. BELL	
Verbal report and judgment of an unstructured stimulus . . . . .	121
BY WILLIAM ROBERT HOOD AND MUZAFAER SHERIF	
A comparative study of the effects of Listica and meprobamate upon motor functioning . . . . .	131
BY M. POWELL LAWTON AND BURTON CAHN	
Racial friction in the deep south . . . . .	139
BY JOHN B. WOLFE AND PAUL HORN	
Stereoscopic depth phenomena with Witte-König effect . . . . .	153
BY J. L. ZAJAC	
Time judgment and schizophrenia: psychophysical method as a relevant contextual factor . . . . .	159
BY FRANCES R. WEBSTER, SANFORD GOLDSTONE, AND WARREN W. WEBB	
Psychological assessment of differential impairment in cerebral organic conditions and in schizophrenics . . . . .	165
BY CHARLES A. KNEHR	
Personal attributes of color and design preferences in clothing fabrics . .	191
BY NORMA H. COMPTON	
Some factors affecting the perception of events as chance determined . .	197
BY SHELDON BLACKMAN	

Sex differences in response set . . . . .	203
BY PHILIP LANGER	
Sensory stimulation and secondary reinforcement . . . . .	209
BY MICHAEL P. ANDRONICO AND DONALD G. FORGAYS	
Contour absence as a critical factor in the inhibition of the decay of a movement aftereffect . . . . .	221
BY IRWIN M. SPIGEL	
Profiles of an adolescent . . . . .	229
BY SAMUEL GOLDMAN	
A comparative analysis of occupational interest scores between gifted and typical 5th grade pupils . . . . .	241
BY RUSSELL N. CASSEL AND GENE HENDSCH	
A study of the effects of mechanically induced tension of the neck muscles on the perception of verticality . . . . .	245
BY CARL W. SCHNEIDER AND S. HOWARD BARTLEY	
Effects of success and failure and the competence of the model on the acquisition and reversal of matching behavior . . . . .	251
BY MILTON E. ROSENBAUM, DOUGLAS K. CHALMERS, AND WILLIAM C. HORNE	
New leader behavior description subscales . . . . .	259
BY RALPH M. STOGDILL, OMAR S. GOODE, AND DAVID R. DAY	
Formation and evaluation of hypotheses in dream interpretation . . . . .	271
BY ERIKA FROMM AND THOMAS M. FRENCH	
Patient wisdom: An anthology of creative insights in psychotherapy . . . . .	285
BY STEWART B. SHAPIRO	
Perceptual functioning in the duckling . . . . .	293
BY NICHOLAS PASTORE	
Reaction-time as a function of stimulus intensity among normal and psychotic subjects . . . . .	299
BY H. E. KING	
Predictability of multidimensional absolute identifications from information transmitted with unidimensional stimuli . . . . .	309
BY GLENN R. HAWKES	
Increased occurrence of EEG alpha during increased attention . . . . .	317
BY THOMAS MULHOLLAND AND SYLVIA RUNNALS	
Personality and attitudinal correlates of response to drug treatment in psychiatric outpatients: I. Theoretical orientation and measuring instruments to be employed . . . . .	331
BY K. RICKELS AND R. W. DOWNING	
Personality and attitudinal correlates of response to drug treatment in psychiatric outpatients: II. Preliminary investigation and implications for further research . . . . .	345
BY R. W. DOWNING AND K. RICKELS	
"Hypnotic behavior" as a function of task motivation . . . . .	363
BY THEODORE XENOPHON BARBER AND DAVID SMITH CALVERLEY	



Predicting organization in changed communication networks . . . .	391
By ARTHUR M. COHEN AND WARREN G. BENNIS	
Measurement of the primary factors of the total personality . . . .	417
By THOMAS N. JENKINS	
Differential impairment in multiple sclerosis . . . . .	443
By CHARLES A. KNEHR	
A discussion of chronogeometry as related to non-stationary time series such as the physiological electrogram . . . . .	453
By SAMUEL J. M. ENGLAND AND BENJAMIN PASAMANICK	
Factor analyses of schizophrenic and organic test data . . . . .	467
By CHARLES A. KNEHR	
The continuity-discontinuity controversy in paired-associate learning . . .	473
By NELSON BUTTERS AND MORTON WIENER	
Consistency between attitudinal affect and spontaneous cognitions . . . .	485
By MILTON J. ROSENBERG AND PHILIP K. OLTMAN	
Stimulus size and Rorschach responses . . . . .	491
By IJA N. KORNER, ROGER B. ALLISON, JR., AND MAX D. ZWANZIGER	

\$20.00 per annum  
\$15.00 per volume  
\$7.50 per half volume

QUARTERLY  
Two volumes per year  
Immediate publication

July, 1962  
Volume 54, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JULY, 1962

(Manuscripts are printed in the order of final acceptance)

The process of free recall: evidence for non-associative factors in acquisition and retention . . . . .	3
BY SOLOMON E. ASCH AND SHELDON M. EBENHOLTZ	
Time judgment and schizophrenia: step interval as a relevant contextual factor . . . . .	33
BY DAVID J. WRIGHT, SANFORD GOLDSTONE, AND WILLIAM K. BOARDMAN	
Freud's conception of the origin of curiosity . . . . .	39
BY JOEL ARONOFF	
Attitudes of a group of nursing students toward a clinical affiliation in a public hospital for the mentally ill . . . . .	47
BY MILDRED E. FLETCHER	
A note on the validity of the Bernberg Human Relations Inventory . . . . .	65
BY DAVID S. GORFEIN AND LARRY M. ANDERSON	

(OVER)

Copyright, 1962, by The Journal Press  
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at  
Provincetown, Mass., under the Act of March 3, 1879  
Second-class postage paid at Provincetown, Mass.



A stimulus-brain feedback system for evaluation of alertness . . . . .	69
BY THOMAS MULHOLLAND AND SYLVIA RUNNALS	
Basic motivation and concept of nursing as chosen profession . . . . .	85
BY EDWARD J. FURST, ALTON W. RAYGOR, AND ALAN P. CROFOOT	
Detection of differences in duration of acoustic and electrical cutaneous stimuli in a vigilance task . . . . .	101
BY M <sub>0</sub> LOEB AND G. R. HAWKES	
Binocular perceptual discriminations of authority <sup>3</sup> and peer group figures among over, under and equal achievers . . . . .	113
BY HAL M. WELLS AND DEANNE M. BELL	
Verbal report and judgment of an unstructured stimulus . . . . .	121
BY WILLIAM ROBERT HOOD AND MUZAFER SHERIF	
A comparative study of the effects of Listica and meprobamate upon motor functioning . . . . .	131
BY M. POWELL LAWTON AND BURTON CAHN	
Racial friction in the deep south . . . . .	139
BY JOHN B. WOLFE AND PAUL HORN	
Stereoscopic depth phenomena with Witte-König effect . . . . .	153
BY J. L. ZAJAC	
Time judgment and schizophrenia: psychophysical method as a relevant contextual factor . . . . .	159
BY FRANCES R. WEBSTER, SANFORD GOLDSTONE, AND WARREN W. WEBB	
Psychological assessment of differential impairment in cerebral organic conditions and in schizophrenics . . . . .	165
BY CHARLES A. KNEHR	
Personal attributes of color and design preferences in clothing fabrics . . . . .	191
BY NORMA H. COMPTON	
Some factors affecting the perception of events as chance determined . . . . .	197
BY SHELDON BLACKMAN	
Sex differences in response set . . . . .	203
BY PHILIP LANGER	
Sensory stimulation and secondary reinforcement . . . . .	209
BY MICHAEL P. ANDRONICO AND DONALD G. FORGAYS	
Contour absence as a critical factor in the inhibition of the decay of a movement aftereffect . . . . .	221
BY IRWIN M. SPIGEL	
Profiles of an adolescent . . . . .	229
BY SAMUEL GOLDMAN	
A comparative analysis of occupational interest scores between gifted and typical 5th grade pupils . . . . .	241
BY RUSSELL N. CASSEL AND GENE HENDSCH	
A study of the effects of mechanically induced tension of the neck muscles on the perception of verticality . . . . .	245
BY CARL W. SCHNEIDER AND S. HOWARD BARTLEY	

## THE PROCESS OF FREE RECALL: EVIDENCE FOR NON-ASSOCIATIVE FACTORS IN ACQUISITION AND RETENTION<sup>\*1</sup>

*Swarthmore College and Connecticut College*

SOLOMON E. ASCH AND SHELDON M. EBENHOLTZ

### A. STATEMENT OF THE PROBLEM

In recall one often proceeds from one datum of past experience to another. It has been an axiom of psychological thinking that consecutive recall is associatively determined, in the sense that specific linkages between contents alone decide the course of recall. Indeed, associationistic doctrines were initially formulated largely in order to account for the "train of ideas." This conception has continued to serve as the explanation of sequential recall to the present day; it is an unquestioned axiom that only sequential associations provide the bridges between the data of past experience, and that recall comes to a halt when the connections give out. We propose to show that under certain conditions consecutive recall is not a function of associative connections between successive contents. For this purpose it becomes necessary to analyze the process of free recall.

It is not difficult to see why this interpretation of sequential recall has escaped challenge. The memory experiment consists of two steps: acquisition and recall. Under the dominance of the association concept it seemed evident that acquisition of successive contents must itself be associative. Consequently the first step of the memory experiment has, from the time of Ebbinghaus, concentrated largely on the acquisition of associations between contents. It is not surprising that the recall subsequently obtained was, or appeared to be, of necessity associative. Given the theoretical starting point, this procedure of investigation seemed adequate to the entire range of recall phenomena. We will see that the consequence was to direct attention to certain properties of recall and to obscure others that are equally relevant to the recall process.

In this inquiry we will rely for evidence on the data of free recall.

\* Received in the Editorial Office on February 6, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This investigation was supported by grants from the Ford Foundation and the National Science Foundation. We wish to thank Mrs. Jean Ebenholtz for her assistance in experimentation and in the analysis of the data.



Although free recall is a well-known phenomenon, and has been frequently employed as a measure of retention, the underlying process has not been investigated. In particular, there has been no systematic analysis of the conditions that determine the order in which materials emerge in free recall (although the effort to explain the latter was the principal aim of associative theories and of the resulting experimental movement). Free recall appeared to raise no new problems for an associative position; it seemed plausible to suppose that the order in which materials appear in free recall is prescribed by the associative connections between them. This is the question that we propose to study.

The present study is the first of a series of investigations of the free recall process. The object is to identify some of the determinants of free recall both during acquisition and in subsequent retention. Since one principal question is whether sequential recall is a function of specific inter-item associations, the procedure of learning will not consist of the formation of such associations, but will rather attempt to eliminate them. To this end the mastery of a set of unconnected items, the order of which is varied on successive trials, is most appropriate. A further question to be studied concerns the effect of interpolation on free recall; accordingly the procedure to be followed will have the design of an RI experiment.

#### B. PROCEDURE

The task consisted of the mastery of unconnected lists of nonsense syllables. The method of learning was designed to reduce or eliminate the likelihood of serial associations; to this end the sequence of items was randomized from trial to trial. Both acquisition and subsequent retention were measured by free recall. The design was essentially that of an RI experiment. There was a total of 24 Ss, divided into three groups of eight—two experimental and one control group.

All experimental Ss (Groups 1 and 2) learned two lists of eight nonsense syllables each to criterion; these followed in immediate succession. The control Ss (Group 3) learned a single list. All Ss received a free recall retention test 20 minutes after mastery of the original list. The experimental Ss were tested for recall of all items from both lists. After recall was completed, all Ss (with the exception of those who recalled the first list completely) proceeded to relearn the first list.

Learning of the lists was by the method of free recall. Each syllable was exposed in the window of a Stoelting memory drum for three seconds, at which time S recited it aloud. Immediately following the presentation and

recitation of the list, *S* was requested to recall, in any order, all the syllables. A period of one minute was allowed for recall. This procedure of alternating learning and test trials was repeated until *S* recalled all eight syllables on a single trial. To reduce the likelihood of inter-item associations, the items were presented in four different orders. Immediately after mastery of the first list (OL), the experimental *Ss* learned the second list (IL) by the procedure just described. In what follows we will refer to the measure of acquisition as "immediate free recall."

The 20-minute interval between the end of learning and the test of retention was occupied with a non-verbal perceptual task. As mentioned above, retention was measured by recall and relearning. At this point the instruction to the two experimental groups differed in one respect. Group 1 was requested to recall the items of OL and IL in any order. *Ss* of Group 2 were directed to recall OL items first; when recall of these was exhausted they were to recall the items of the interpolated list (IL). The object of the latter variation was to observe the effect of directed instructions on the course of recall. Experimental *Ss* had a maximum of two minutes for recall, control *Ss* a maximum of one minute. All *Ss* were directed to guess freely. The test of relearning OL followed immediately upon the conclusion of the recall test, the procedure being that employed in original learning. Recall and relearning were also measured by the method of free recall.

There were two lists (A and B) of comparable association values. Four syllables of each list were of 47 per cent association value (9), and four of 53 per cent value. Items were chosen so as to maintain as low a level as possible of inter- and intra-list similarity. One half of the *Ss* in each experimental group learned one of these lists as OL and the other as IL; the two lists were also evenly divided among control *Ss*. This counterbalancing procedure was observed in order to avoid effects due to unequal difficulty of the lists; for this we could not rely entirely on identical association values.

*Ss* were assigned to Groups 1 and 3 alternately, according to the order of their appearance in the laboratory; *Ss* of Group 2 were run subsequently. *Ss* were men and women Swarthmore undergraduates.

### C. RESULTS

#### 1. *Learning and Retention*

##### *a. Immediate free recall.*

(1). *Learning of OL.* One-half of *Ss* in each group learned List A, the other half learned List B as their original list. Although the lists were of



identical association values, it was necessary to check their relative ease of learning. The twelve *Ss* who learned List A required 3.9 trials to reach criterion; the mean number of trials to reach criterion with List B was 5.2. By the Wilcoxon unpaired replicates test,  $p > .05$ . There was thus a tendency for List A to be learned somewhat more rapidly, although the difference did not reach significance. Accordingly, the analyses to follow are based upon the combined data of Lists A and B.

Groups 1, 2 and 3 reached criterion in learning their original lists in 3.9, 5.5 and 4.4 trials, respectively. Applying the Kruskal-Wallis one-way analysis of variance, the obtained statistic ( $H$ ) was .047, which for 2 *df* was not significant ( $p > .7$ ). We conclude that the groups did not differ in rate of learning OL.

(2). *Learning of IL*. Mean trials to reach criterion in learning IL were 2.6 and 3.5 for Groups 1 and 2, respectively. By the Wilcoxon test, the difference was not significant ( $p > .05$ ).<sup>2</sup> Both groups mastered<sup>a</sup> IL more rapidly than OL; the differences were not, however, significant.

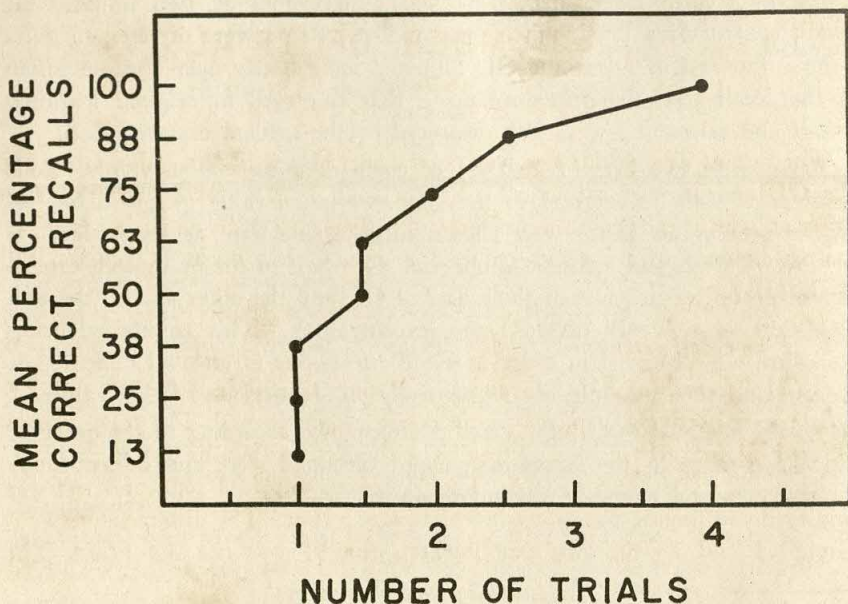


FIGURE 1  
RATE OF OL ACQUISITION: GROUP 1

<sup>2</sup> The difference between the means of Groups 1 and 2 was due to the inordinately high score of one *S*.

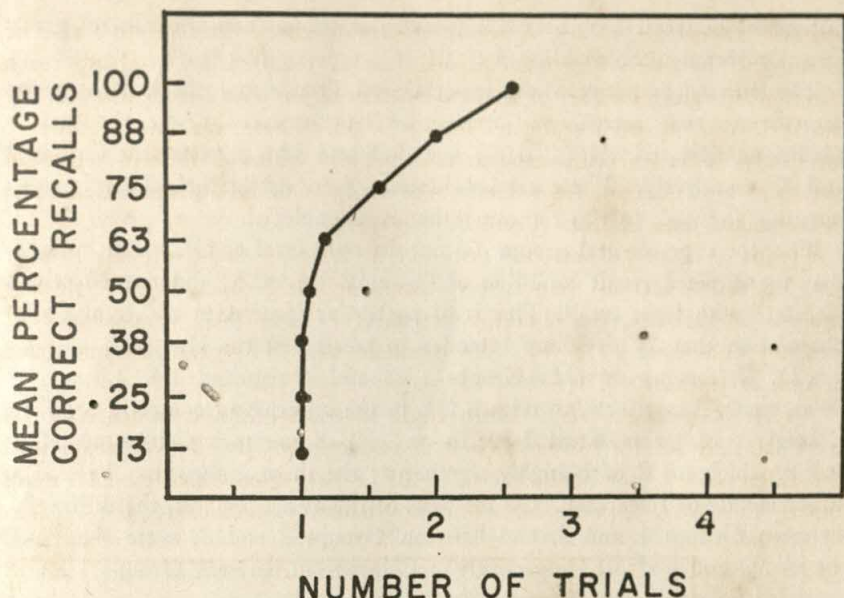


FIGURE 2  
RATE OF IL ACQUISITION: GROUP 1

The rate of item acquisition was rapid under all conditions. Illustrative are the OL and IL acquisition curves for Group 1 (Figures 1 and 2). The criterion was in each case reached after less than four trials on the average; to master one-half of the series required a mean of 1.5 trials (OL) and 1.1 trials (IL).

*b. Free recall retention.*

(1). *Recall of OL.* Groups 1, 2 and 3 recalled 4.8, 5.1 and 7.6 syllables, respectively, of a maximum possible eight items.<sup>3</sup> The three groups differed significantly in level of recall (by the Kruskal-Wallis test  $H = 12.51$ , which for 2 df yields  $p < .01$ ). The Wilcoxon unpaired replicates test was employed to compare the groups taken two at a time. The difference between Groups 1 and 2—the two experimental groups—was not significant. The

<sup>3</sup> The measure employed throughout was correct recalls; results are reported in terms of number of correct recalls or number of trials to reach criterion. We do not report the occurrence of errors, since these were not fixed by the same determinants for all *Ss*. One *S* may give a fair number or a few errors, others none, depending on circumstances over which there was no experimental control.



differences between each of the experimental groups and the control group were significant, each yielding  $p < .01$ .

The differences between the experimental groups and the control group demonstrate that retroactive interference had occurred under the present conditions. Relative RI (C-E/C) was 36.8 and 32.8 per cent for Groups 1 and 2, respectively. These are substantial effects, given the relatively short retention interval (of 20 minutes) that was employed.

That the experimental groups did not differ in level of OL recall indicates that the directed recall condition of Group 2 (cf. p. 5) did not effectively interfere with item recall. This is of particular interest in the light of the observation that Ss of Group 1 tended to recall first the IL items.

(2). *Relearning of OL.* Groups 1, 2 and 3 required 1.4, 1.8 and .4 mean trials, respectively, to relearn OL to the criterion of complete recall on a single trial. The Kruskal-Wallis analysis of variance yielded an  $H$  of 9.398, which for 2 df is highly significant; the three groups may be said to differ in rate of relearning. On the basis of the Wilcoxon test, the differences between Groups 1 and 3, and between Groups 2 and 3, were significant ( $p < .02$  and  $p < .01$ , respectively). Comparison between Groups 1 and 2 revealed no significant difference ( $p > .05$ ).

The analysis of the relearning scores fully supports the results from the test of recall, the differences between each of the experimental groups and the control group indicating significant amounts of RI.

(3). *Recall of IL.* Groups 1, 2 and 3 recalled 6.5, 5.8 and 7.6<sup>4</sup> items, respectively. The Kruskal-Wallis test yielded an  $H$  of 6.8886, which for 2 df is significant ( $p < .05$ ).

Although there was a tendency for Group 2 to recall somewhat less than Group 1, the Wilcoxon test yielded no significant difference ( $p > .05$ ). Both experimental groups, however, recalled significantly fewer syllables than the control group (Group 3), both differences indicating significance at better than the .05 level. Since it is doubtful that the difference in the retention interval (see footnote 4) between the experimental and control groups favored the control group, it may be concluded that proactive interference has been demonstrated for the retention of items. Relative PI (C-E/C) amounted to 14.5 per cent for Group 1, and 23.7 per cent for Group 2.

<sup>4</sup> The control group (Group 3) is the same as that used above in the evaluation of RI. Since the retention test occurred twenty minutes after learning of OL was completed, the time between learning and test for IL was approximately sixteen minutes (IL requiring an average of four minutes to learn). The learning-retention intervals thus differ slightly in the experimental and control groups; however, the four minute difference is assumed to be of little consequence.

(4). *Comparison of Groups 1 and 2.* The preferred mode of recall of Group 1, which was free to give items from the two lists in any order, was to give IL items first. Group 2 was directed to recall OL items first; the effect was to raise slightly the recall of OL and to depress somewhat the recall of IL; neither of these effects reached significance. The induced changes in the order of recall did not appreciably alter the levels of performance. (For a further account of the effect of directed recall see pp. 18-19).

## 2. *Sequential Dependencies in Immediate Free Recall*

A theory of the free recall process requires an explanation of the circumstances responsible for the order in which items become accessible to recall. It is in particular necessary to establish whether the order of free recall was associatively determined. It is conceivable that reference to the general situation initiates the first recall, and that subsequent recall is a function of specific associations. The present experiment purports to deal with the learning and retention of individual items in the absence of specific inter-item associations; the procedure of learning aimed to eliminate the latter. It must, however, be determined independently whether such associations nevertheless occurred, and whether they controlled the order of recall.

The analyses that follow pertain to free recall at the point of acquisition; they are based on the data obtained from all 24 Ss during the *learning* of the first list (OL). The principal question is whether the order of recall corresponds to the order of items in the learning series.

a. *The Influence of Adjacent Items in Learning upon Paired Recall.* To what extent were adjacent items in free recall identical with adjacent items as they appeared on the immediately preceding learning trial? Recalls of such sequences may be regarded as an index of specific inter-item associations.

For each S, the frequency of pairs of successive items which appeared in identical sequences in both recall and learning (recitation) was calculated, as well as the total number of recalled pairings (regardless of whether or not the pair occurred in the recitation sequence). The ratio of repeated to total pairings was then determined. As an illustration of the procedure we present the data of one S on his criterion trial. The temporal order of items in their recitation sequence may be represented by the numbers one through eight. The temporal order of the same syllables as they were recalled (immediately after recitation) was 1, 2, 6, 7, 5, 4, 3, and 8. The total number of pairs recalled was seven; of these, two pairs (syllables 1-2 and 6-7) occurred in the same *pair-order* in both learning and recall. Hence



the *S* was scored with two repeated pairs out of seven. This procedure was followed for all trials, and then summed to give each *S* a single ratio. The average ratio for all twenty-four *Ss* was .18; thus, on the average, only 18 per cent of all pairs given in recall duplicated the sequence in which they were originally recited. We also computed the ratio on the basis of the first recall test only; the obtained value was .17. We conclude that the sequence in which items occur during learning had a limited effect upon the order of immediate free recall.<sup>5</sup>

*b. The Effect of Serial Order in Learning Upon Order of Recall.* The following analysis takes into account the relation between the order of all items recalled on a given test and the order they had in the immediately preceding learning (recitation) series. It does not limit itself to adjacent order.

The index employed was that of "relative correct order." For each *S*, on each trial, the total number of pairs recalled, and the number of pairs in correct order relative to the immediately preceding learning series, was determined. To illustrate, on a given recall test one *S* recalled items 1, 2, 8, 6, and 3 of the learning series, in that order. The total number of pairs recalled was four; two of these (1-2 and 2-8) preserved the relative order of the series. The total number of pairs on all trials, for each *S*, was determined, and also the total of correct relative pairs; the ratio of the latter to the former was the index of correct relative recall. This measure includes adjacent as well as remote pairs in correct order.

The mean ratio for all *Ss* was .51. Calculations made on the basis of the criterion trial alone yielded a mean value of .56; the index based on the first trial only was .54. We assume that on a chance basis relative correct recalls would occur with a frequency approximating .5, which corresponds closely to that obtained. There was no significant tendency to preserve in recall the order of the items given during learning.

*c. Repeated Recall of Pairs.* It is possible that *Ss* formed specific associations between certain items, whether or not these corresponded to the order of learning, and that these were repeatedly recalled. Accordingly, we will now examine the extent of repeated recalls of identical pairs on successive recall tests.

For each *S*, paired recalls were obtained on each trial simply by counting two consecutive syllables as a pair. For example, if on a given trial four

---

<sup>5</sup> The present index probably exaggerates the estimate of inter-item associations. Recall of pairs identical with those of the learning series may also be a function of the relative availability of the items, and of random factors.

syllables were recalled in the order a-b-c-d, this would result in a tally of three pairs (a-b, b-c, c-d). The total number of different pairs recalled on all trials combined, excluding the criterion trial, was determined for each S.<sup>6</sup> We then counted the number of pairs recalled more than once by each S. The ratio of pairs recalled more than once to the total number of different pairs of a given S was taken as his index. The mean ratio for the 24 Ss combined was .23; thus, on the average, 23 per cent of all pairs tended to be recalled more than once. Subsequent analysis revealed that of the 23 per cent repeated, 71 per cent were repeated no more than twice. The apparent lack of repetition in recall of pairs yields no evidence for the substantial formation of inter-item associations.

The following is a somewhat different analysis of the data. If Ss had recalled each pair in the same sequence on all recall trials, each would have shown a total of seven different pairs. If Ss had never repeated a pair, the mean number of different pairs per subject would have been 24.6 (the latter figure representing the mean number of recalled pairs). The mean number of different pairs recalled was 18.9. The difference between 18.9 and 24.6 is indicative of a rather limited tendency toward repetition of pairs in recall.

To what extent were the repeated pairs duplicates of the recitation sequence? There was a total of 95 repeated pairs, and these were given in 230 instances. (1) Of the total of 95 pairs, 43 (45.3 per cent) were identical with pairs in the immediately preceding recitation sequence. These occurred in 53 out of a total of 230 instances (23.0 per cent). (2) Of the 95 repeated pairs, 37 (38.9 per cent) had not occurred in any recitation sequence. The number of such instances was 88 out of a total of 230 (38.3 per cent). (3) Fifteen out of the total of 95 repeated pairs (15.8 per cent) had occurred at some point prior to the immediately preceding recitation sequence. There was thus a total of 58 (or 61.1 per cent) repeated pairs that appeared at some point during recitation. No possible associative effect is plausible in 38.9 per cent of the repeated pairs. The remaining instances may reflect recitation order effects, but the results under (3) probably include random recitation-recall matches. Again it appears that associative effects are limited.

d. *The Influence of Primacy and Recency upon Order of Recall.* The preceding analyses have identified a number of conditions that fail to account more than slightly for the order of recall. As a final attempt to isolate a characteristic determining property, we examined the recall of items at the extremes of the series. The measure was the frequency with which the first

<sup>6</sup> The criterion trial was excluded, since pairs appearing first at this point had no opportunity to be repeated.



and last items in the learning (and recitation) sequence were recalled first on the immediately following recall test. For each *S* two proportions were determined: (a) the number of recall trials in which the first recitation item was recalled first (primacy), divided by the total number of recall trials; and (b) the number of recall trials in which the last (*viz.*, the eighth) item in the recitation series was given first in recall, divided by the number of trials (recency).

The mean ratio for 24 *Ss* combined was 32 per cent for primacy, and 27 per cent for recency. Taken together, the first and last items to be recited were the first to be recalled in 59 per cent of all trials; the items occupying the six remaining positions were recalled first in only 41 per cent of all trials. If items in the first two recitation positions are chosen to evaluate the primacy effect, and items in positions seven and eight are taken as an index of recency, the former were recalled first on 35 per cent of all trials and the latter were recalled first on 38 per cent of all trials. Taken together, the first and last pairs in the recitation sequence were the first to be recalled in 73 per cent of all trials; the items occupying the remaining four positions occurred first in recall only 27 per cent of the time.

TABLE 1  
ORDER OF RECALL AS A FUNCTION OF SERIAL POSITION (LO): ALL RECALL TESTS

Serial position (LO)	1	2	3	4	5	6	7	8
Frequency of first and second recalls	50	28	17	17	18	19	27	44

The following is a more exhaustive analysis of the data. It lists frequency of recall and recall order (RO) as a function of learning order (LO), or of position in the recitation sequence. Included are first and second recalls only. The results appear in Table 1. Items in positions 1, 2, 7 and 8 gave 149 recalls, those in the middle positions (3, 4, 5 and 6) contributed 71 recalls. The respective percentage values, 67.7 and 32.3 per cent, are quite similar to the preceding results.

We also computed the primacy-recency effects for the first recall trial alone; the results appear in Table 2. The sixteen *Ss* who contributed scores

TABLE 2  
ORDER OF RECALL AS A FUNCTION OF SERIAL POSITION (LO): FIRST RECALL TEST

Serial position (LO)	1	2	3	4	5	6	7	8
Frequency of first and second recalls	5	4	1	2	3	3	5	8

to the present analysis gave a total of 31 recalls on the first test. Of these, 22 (70.9 per cent) came from positions 1, 2, 7 and 8; nine recalls (29.1 per cent) came from positions 3, 4, 5 and 6. These results closely support those based on all recall trials.

Items in the first two and last two positions were likely to be recalled prior to items in the central positions. Primacy and recency effects thus account in large part for the order of recall in free recall.

### 3. *Determinants of Immediate Free Recall*

The preceding analyses of the sequential order of recall were of necessity limited to items that figure in recall regardless of their frequency. The differential availability of items constitutes another fundamental aspect of the free recall process. In what follows we will examine some of the determinants of item availability.

*a. Primacy-Recency Effects.* Primacy and recency are, we saw, a significant condition of order of recall. Do they also determine the occurrence of recall? Since the sequence of items in recitation was varied on successive trials, the most pertinent data are recalls obtained immediately following the first recitation trial. Figure 3, which plots the frequency of correct recalls for all 24 Ss as a function of serial position during learning of the original list, shows that first and last items were indeed more frequently recalled than those in other positions. (It should be mentioned, however, that since only four recitation sequences were employed during learning, each item did not appear in each of the serial positions an equal number of times.)

We conclude that primacy and recency are conditions of item mastery in immediate free recall. The data do not, however, present convincing evidence that items beyond the extremes of the series differed in rate of mastery. In the absence of more adequate evidence one may therefore not conclude that the serial position curve for items in free recall is similar to the bow-shaped curve for serial learning, or that the underlying processes are identical in the two cases. The possibility must be considered that the effect here reported pertains exclusively to the extremes of the series.

### 4. *Long-Term Retention*

The following analyses are based on the final recall test that came twenty minutes after learning.

*a. Role of Alphabetical Order.* It appeared necessary to check the possibility that Ss employed the alphabetical order of the initial letters of



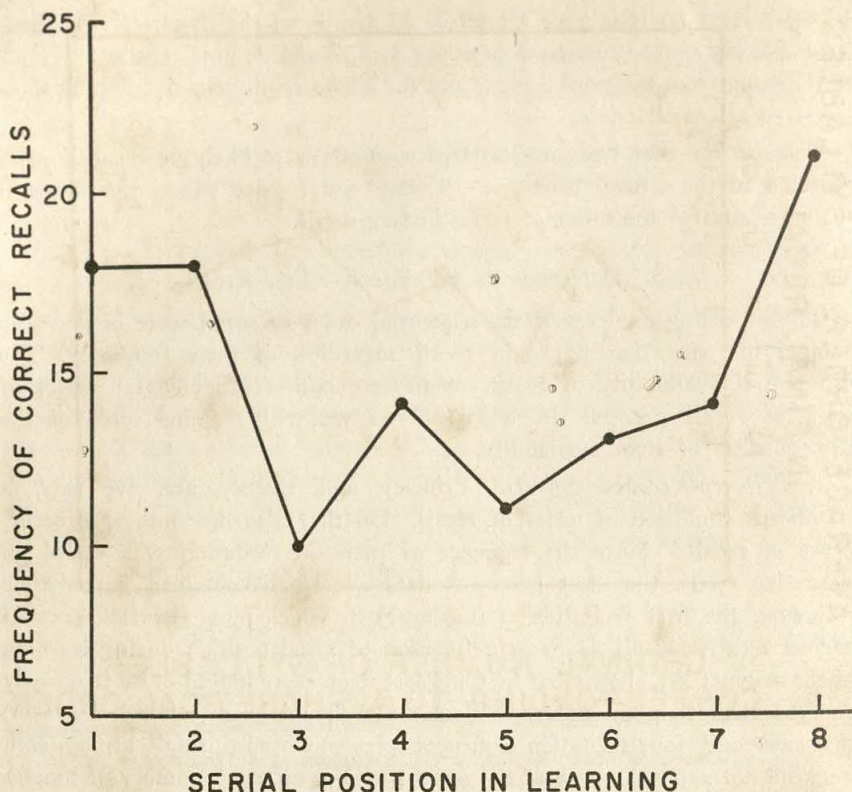


FIGURE 3  
SERIAL POSITION AND FREQUENCY OF RECALL

the syllables as an aid to recall. Such a procedure would serve to transform the free recall process into one of aided recall.

For this analysis, the data of recall were those of the control group (Group 3). These were uncomplicated by the interpolation of another series; also, recall of this group was near maximum as six of eight *Ss* recalled perfectly; the others recalled six and seven items (out of eight), respectively.

The syllables of the series were placed in alphabetical order, and the order of recall of each syllable was recorded for each *S*. The mean order of recall for each syllable was then computed by averaging across all *Ss*. The results are plotted in Figure 4.

A close relation between alphabetical order and order of recall would

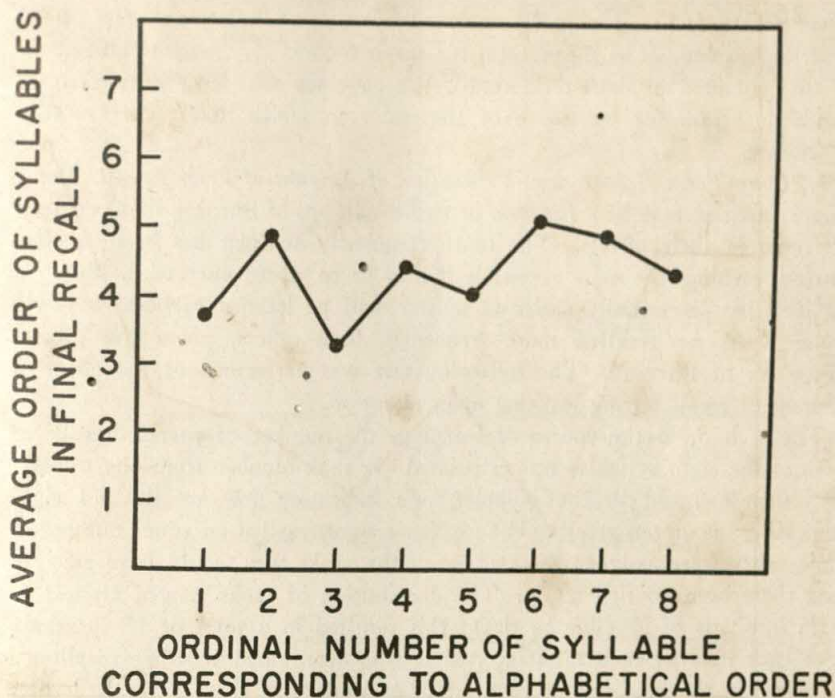


FIGURE 4

ORDER OF FINAL RECALL OF OL SYLLABLES AS A FUNCTION  
OF ALPHABETICAL ORDER: GROUP 3

give a 45 degree slope, intersecting the axes at the origin. The obtained function shows little or no relation between the two measures. This finding is supported by the reports of most *Ss* that they did not rely on the alphabetical sequence in recall.<sup>7</sup> The same analysis was applied to OL recalls of the experimental groups, and with the same results; the latter are, however, less conclusive in view of the decrement of recall and of the requirement to recall items from two lists.

*b. Effect of Order of Preceding Recall.* Is the order of recall in long-term retention a function of the order of preceding recall? The order of recall was compared with the order obtained on the criterion trial. Only control *Ss* were included in this analysis; one of these was eliminated from the calculation, since he had employed the alphabet as a basis for ordering his

<sup>7</sup> One *S* ordered his recalls alphabetically. The data for this subject are included in Figure 4.



recall (footnote 7). The index employed was that of "relative correct order" described earlier. The mean ratio for seven *Ss* was .57, a value quite close to that obtained in short-term recall. We conclude that long-term retention yields no significant indication of the order in which items were recalled previously.

*c. Long-Term Retention as a Function of Amount of Prior Recall.* Long-range retention may be a function of two conditions of learning that generally converge in their effects. The more frequently an item has been recalled during learning, the more accessible it may be to subsequent recall. Further, an item that is initially easier to master will be learned earliest, and will consequently be recalled more frequently than others, given the present procedure of learning. The following test was performed of the effect of amount of recall during learning upon retention.

For each *S*, in the course of learning, the number of correct recalls of each of the eight syllables was calculated. A rank number from one to eight was then assigned to the syllables, one indicating few recalls and eight indicating maximum recalls. Where items were recalled an equal number of times, they were assigned the average of the ranks they would have received had there been no tied scores. The distribution of ranks ranged from 1 to 8 in intervals of .5 (due to ties); this resulted in a total of 15 intervals. For each item at each rank interval it was determined whether recall had occurred in the final test of recall. For example, if a given *S* recalled three items (during learning) at rank 1.5, it was noted whether or not these three items were retained on the final retention test. Items that occurred at the same rank were pooled for all *Ss* in the control and experimental groups, respectively, and their corresponding frequency of recall during the retention test was determined. For purposes of plotting, the intervals were combined to form three groups, i.e., ranks 1.0-3.0, 3.5-5.5 and 6.0-8.0. For each of these categories the percentage of items recalled in the final retention test was determined. The plot of the data for original and interpolated lists appears in Figure 5.

(1) There was no recall gradient in the control group. Its absence does not, however, imply a failure of the relation in question. Since recall was in this condition nearly complete (95 per cent of maximum), it rather suggests that item availability reaches optimum effectiveness at relatively low levels of practice. To estimate the effects of prior recall on subsequent retention, it is necessary to study conditions that give an appreciable recall decrement. For this reason the experimental groups are of particular relevance. Interpolation in these groups produced a substantial decrement;

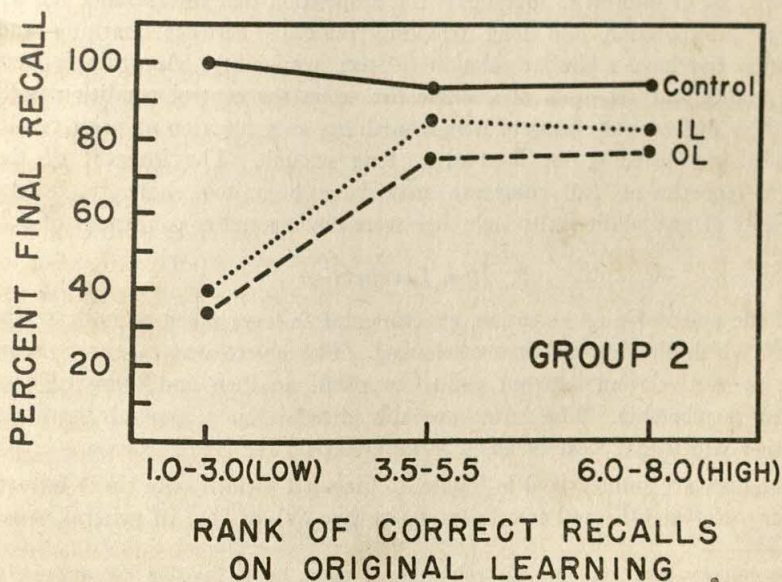
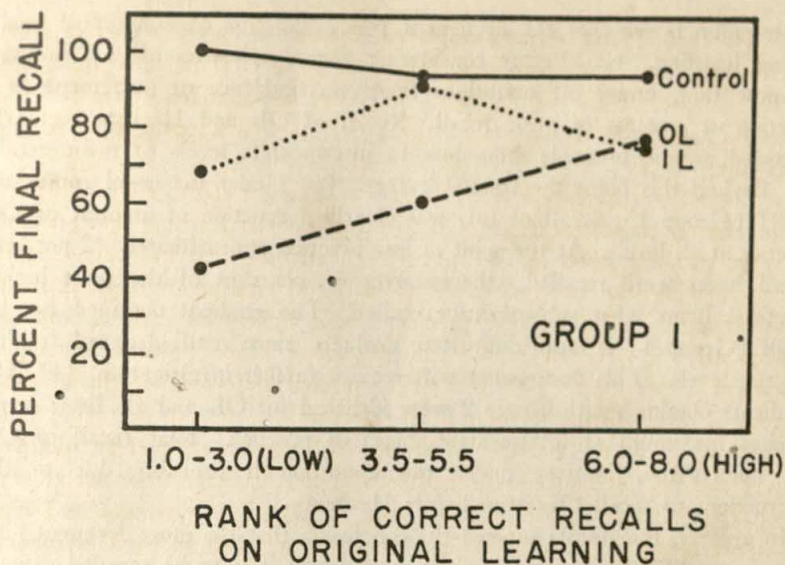


FIGURE 5  
FINAL RECALL AS A FUNCTION OF AMOUNT OF ORIGINAL RECALL PRACTICE



the question is whether the decrement was a function of amount of recall during learning. (2) Before considering the complexities of the findings, we note that, under all conditions there was evidence of decrement as a function of amount of prior recall. Recall of OL and IL lists regularly increased as one proceeds from low to intermediate levels of prior recall. (3) Beyond this point the results diverge. (a) Under the usual conditions of RI (Group 1), recall of OL was clearly a function of amount of OL practice at all levels. At the point of low practice, approximately 42 per cent of all items were recalled, whereas over 77 per cent of the most highly practiced items were subsequently recalled. The gradient obtained for IL recall (Group 1) is more difficult to evaluate, since recall dropped at high practice levels. This discrepancy will require further investigation. (b) The gradients obtained with Group 2 were identical for OL and IL lists; recall reached maximum at intermediate stages of practice. That recall of OL did not further increase under this condition is probably due to the instructions to recall OL items before IL items.

In general, the results support the conclusion that the more frequently an item is recalled during learning the more likely it is to be recalled after a retention interval.<sup>8</sup>

It may be of interest to investigate the proposition that interference factors due to interpolation and long temporal intervals between learning and retention test have a similar relation to item availability. Accordingly, one may predict that retention of a single list, as in the control condition, will also show differential effects of item availability as a function of prior recall, provided the learning-test interval is long enough. The interval of the present experiment (20 minutes) may have been too short to exhibit gradients of availability, although they were obvious under conditions of RI.

### 5. *Item Localization*

At the point of final retention, experimental Ss were asked to indicate the list to which the recalled items belonged. The object was to throw some light on the relation between ability to recall an item and knowledge of its list membership. The latter probably depends on a general localizing function which may bear in an important way on the recall process.

The data are summarized in Table 3. Group 1 yielded only three failures (by two Ss) in OL, and two failures (by two Ss) in IL. In general, when

<sup>8</sup> It follows that rate of relearning, too, should be a function of amount of earlier recall. Relearning in the present experiment was too rapid to permit an analysis of this relation.

TABLE 3  
LOCALIZATION OF LIST MEMBERSHIP OF ITEMS

	Mean items recalled	OL Mean items correctly localized	Per cent of correct locali- zations	Mean items <sup>a</sup> recalled	IL Mean items correctly localized	Per cent of correct locali- zations
Group 1	4.8	4.4	91.7	6.5	6.2	95.4
Group 2	5.1	3.8	74.5	5.8	5.1	88.0

an item was recalled there was a high probability that it would be correctly localized. The conditions of Group 2 produced some disturbance of the localization process. Six of eight Ss contributed a total of 11 incorrect identifications in OL, all of which were incorrectly referred to IL; of five failures to localize in IL, two were incorrect assignments, and three were instances of uncertainty. In the light of the results of Group 1, it is apparent that the directed recall instructions tended to interfere with the localization function in OL, although they did not adversely affect recall of the items. The relation between recall and localization requires further investigation.

#### D. DISCUSSION

##### 1. *Consecutive Recall and Association*

Under the conditions of this investigation successive recall was not substantially a function of specific associations between items. (a) The order in which items were produced in short-term recall (which registered the course of acquisition) did not correspond notably to the order of earlier experience. Further, acquisition was rapid. Although this investigation does not provide a specification of what constitutes rapid learning, it is doubtful whether the rate of acquisition could be further accelerated by the addition of inter-item associations.<sup>9</sup> (b) One must conclude that long-term retention, too, was not a function of inter-item association. (c) It also follows that loss of retention was not referable to the weakening or loss of associative connections between items. These findings contradict the assumption that associative linkages between items alone determine the order in which they emerge in recall, and that retention loss is generally associative.<sup>10</sup>

<sup>9</sup> We will subsequently report an investigation that compared the free recall of identical contents under two conditions: when they were associated with each other, and when associative connections were excluded. Recall was the same under the two conditions.

<sup>10</sup> The objection may be raised that the learning procedure here employed produced an equal number of associations between all items, and that the associations



Under other conditions free recall will tend to follow the course of earlier experience. Thus one often (but not necessarily) retells a story in the order in which one heard it; the recalled sequence will tend to parallel that of original experience, most likely because both sequences correspond to the intelligible relations between the events referred to. Free recall may also preserve the order of earlier experience in the absence of meaningful organization. For example, a serially learned list of nonsense syllables may subsequently be recalled in the order that has become habitual; in the absence of a competing organization, later retention may follow the lines of earlier retention. The findings here reported are clearly specific to a procedure that eliminates or reduces inter-item associations. However, to stress that the effects were obtained under particular conditions would be to miss the significance of this investigation. The fact of consequence is that there are conditions capable of producing adequate acquisition and recall that are not under the control of inter-item associations.

## 2. *The Concept of Availability*

A theory of recall must account for differential recall of data as well as for order of recall. The prevailing interpretation of differential recall is also strictly associative, the assumption being that what is well recalled is what has been well associated. Since differential recall under the present conditions occurred largely in the absence of specific inter-item associations, the latter cannot account for differences in recall of items.

No assumption has been more deeply entrenched in psychological thought than that which derives the phenomena of recall from association. Its current expression is in terms of S-R conceptualization. According to the latter, each instance of recall is a "response"; and a response is defined as that which is elicited by a stimulus with which it has been associated. This analysis is responsible for the interpretation that the ease of recall of a datum is a function of the number of associative connections previously formed to it; it is at the basis of the attempts of recent decades to scale the ease of "learning" of nonsense syllables on the basis of their "association value" (9, 10, 11, 12, 15). It is also the reason that studies of differential recall of items employ the procedure of the association experiment, equating recall with "response availability" or with "response integration" (23).

---

were of equal strength. On this basis one might attempt to explain the relative absence of a relation between the order of recall and order of recitation. However, the assumption that all items were equally associated cannot account for the obtained effects of primacy and recency, nor for the finding that retention increased as a function of amount of prior (item) recall.

These efforts are understandable only as a consequence of the absorption of recall into association.

This investigation points to a fundamentally different theory of the recall process. The concept we will employ is that of availability, which acquires a new and systematic importance in the light of the preceding findings. Availability refers to the accessibility of an item to recall; the index of availability is the ease or frequency with which a datum of past experience can be produced. Central to our conception is the definition of availability as a phenomenon of recall, distinct from association between items; availability refers to a property of memory traces, with respect to which they differ. We propose that the laws of availability are not laws of association. The relation between these is a fundamental problem that can best be clarified if the respective processes are not confused. In a recent investigation of "backward association" (1) we provided support for the distinction in question. The widely held view that backward associations are weaker than those in the forward direction was, we showed, a consequence of the failure to distinguish between a property of associations and the independent process of recall, or of attributing to association what belongs to recall. It was possible to show that it is erroneous to conclude from the unavailability of an associated item to the absence of the association, or to its low strength. The present investigation further demonstrates that differential recall is not a function of specific inter-item associations.

The conceptual distinction between association and recall sets before us the task of specifying the conditions of availability. It further suggests that this may be done most adequately without the introduction of specific associative bonds between items, and that the procedure of free recall is best suited to the purpose. We will now summarize several conclusions about the conditions of availability deriving from this investigation; shortly we will examine the question in the light of other studies.

*a. Frequency is a determinant of availability.* This familiar finding must now be considered as an effect independent of specific inter-item associations: we found that availability increased as a function of frequency, in the absence of such associative formations. The relevant evidence from the present investigation may be summarized as follows: (a) Recall of items increased on successive trials; (b) Long-range retention was a function of frequency of prior recall.

*b. Primacy and recency are conditions of short-range free recall.* Two effects were observed, which must be distinguished and related to each other: serial primacy and recency determined both the probability of recall and order



of recall. The items recalled more often were also recalled earlier than others. There are thus two correlated indices of availability: probability and latency of recall. We call attention to the fact that these have hitherto been considered indices of association or "habit-strength." These effects of primacy and recency cannot be derived from inter-item associations.

### 3. *Availability and Retroactive Interference (RI)*

Because of its special place in the theory of recall, we consider separately a third condition of availability—that of interpolation. We obtained substantial RI decrements, as measured both by recall and relearning in the absence of specific inter-item associations. These findings raise a new question concerning the interpretation of interference phenomena.

Current theories of interference are, like those of acquisition and recall, purely associative. This is evident first in the procedures of studying interference effects; the classical paradigms are associative in design (e.g., the A-B, A-K paradigm). Second, the theories in question deal exclusively with interference effects upon associative bonds.

The first of the two principal interpretations of interference is the competition of response theory stated by McGeoch (13). It asserts that if two associations are attached to the same stimulus (A-B and A-K) they compete at recall, and that the stronger will block the weaker. This principle is not a tenable explanation of the present results. Competition of response refers specifically to associative connections between units; these, as we saw, were of minor importance in this investigation. Further, the procedure we followed did not induce competition between units. Unlike the procedure frequently followed in RI experiments, we did not require *S* to make a choice, at the point of recall, between alternatives that might disturb each other; the test of recall referred equally to the interfering and interfered-with lists. (For a further discussion of this point see pp. 25-26.)

The second major mechanism proposed in explanation of RI is the unlearning hypothesis of Melton and Irwin (14). It asserts that, in addition to competition of response, interpolated learning of associations produces an unlearning of preceding associations. Since few associations were formed in the present investigation, there could hardly be associative unlearning. We have thus demonstrated RI effects when the conditions were not given either for competition of responses or for unlearning. Despite the absence of these conditions, RI was quite substantial and not unlike the levels frequently reported under associative conditions.

What is the bearing of the present findings on the interpretation of

interference effects? It may be helpful to state first the logical possibilities. (a) One may first ask whether the effects of interpolation here reported are reducible to associative interference as currently described. (b) One might maintain that there are two distinct sources of interference—associative and item interference; that the former is valid for the associative conditions on which investigation has primarily focussed, and that the latter is relevant to conditions that exclude inter-item associations. (c) One must also consider the possibility that, generally, *interference effects do not act upon the associative bond*, or by means of the associative bond; that they are purely effects on availability of items. If so, the concepts of competition of response and of unlearning would become unnecessary to the interpretation of RI effects.<sup>11</sup>

a. In an earlier investigation that reported results similar to ours, Postman (16) offered an explanation of the resulting interference in associative terms. Employing an RI design with lists of unconnected syllables, and with the procedure of free recall, Postman obtained a significant decrement of recall for the first (OL) list. He referred the decrement to competition of responses and to unlearning. It should be pointed out that Postman did not consider the possibility that acquisition under conditions of free recall is not associative. The retention decrement was, he proposed, due in part to "competition between original and interpolated responses" (16, p. 168). To clarify the grounds of this statement, we need to describe the procedure of this experiment. Following a usual practice, Postman required Ss, at the time of test, to recall only the OL items. As evidence for competition of responses, he cited the occurrence of intrusions, or of IL items, during the test of recall. This inference seems to us questionable. In the absence of more direct evidence it is erroneous to equate intrusions during recall with response competition. Such intrusions may signify rather that Ss erred in assigning the items they recalled to their proper list. The procedure of testing contains the implicit requirement that S must differentiate between the lists in question. It is conceivable, however, that S recalls the OL items but not their list memberships. A similar comment applies to the hypothesis of unlearning that Postman also invoked in interpretation of his findings. This concept initially had a clear sense; it referred to the disruption of an associative connection. In the present context, the data tell us only that there was loss of retention. To

<sup>11</sup> Here we may also note that the application of the concepts of unlearning and of competition of responses becomes somewhat tenuous when one attempts to explain RI effects obtained under A-B, C-D conditions, or when OL and IL are dissimilar in content.



equate loss of retention with the unlearning of associative bonds is to confuse a phenomenon with its explanation; the two are far from synonymous, and they may be entirely unrelated. In the light of our findings, it seems necessary to conclude that the interference in question was not an effect upon associative bonds.

*b.* The second alternative posits two sources of interference, the relative contributions of which may differ depending upon the conditions of learning. A decision concerning this formulation must await further evidence. It should be stated, though, that the availability interpretation of interference cannot be restricted to the conditions here investigated. To the extent that loss of availability is responsible for RI effects, it must also contribute in full measure to RI under associative conditions. On the other hand, associative interference cannot, for the reasons stated above, account for the effect of interpolation upon unassociated items.

*c.* It remains to be considered whether interference effects are wholly due to loss of item availability. In view of the vast effort that has gone into the associative study of RI, it would be premature to draw a definitive conclusion on the basis of a single investigation, all the more so since it did not directly analyze interference effects in the associative situation. It would accordingly be in order to systematically compare RI (and PI) effects under associative and free recall conditions. By varying the known determinants of interference, such as degree of similarity and amounts of OL and IL learning, it should be possible to establish whether the two procedures yield parallel functions, and thus to decide whether a loss-of-availability interpretation of interference accounts completely for interference effects.<sup>12</sup>

Returning to the present investigation, we have demonstrated that interpolation reduces availability. The effect is transitory, as is evident in rapid relearning. The mode of action of interpolation is at present unknown. The effect may be in part that of recency, which was also shown to be an important condition of intra-list retention. It would be of value to determine whether RI for items is particularly strong when IL immediately follows OL. Investigation is also needed of the relation between inter- and intra-

---

<sup>12</sup> One investigation that we have recently completed (still to be published), which employed the A-B, A-K design, demonstrated that the obtained RI decrement could be wholly accounted for in terms of loss of availability.

Although an examination of the literature must be reserved for the future, we refer the reader to the independent studies of Ceraso (4) and Runquist (19); both investigators found no evidence of *specific* interference. These findings constitute, we believe, a significant challenge to an associative interference theory.

list interference, and of the proposition that interpolation and long temporal intervals between learning and test of recall have a similar relation to availability.

#### 4. Context Localization

This investigation provided certain preliminary findings concerning context localization (Table 3). Given the RI design, the materials belonged (for the experimental groups) to one of two contexts, which they were asked to identify concurrently with recall. The procedure has been previously employed by Barnes and Underwood (2); its rationale follows closely from Underwood's (22) concept of differentiation. We found that localization of context was highly accurate under usual conditions (Group 1), but was measurably disturbed with instructions directing the order of recall (Group 2); yet even in the latter instance assignment to context was predominantly accurate. Further investigation is needed of the process of localization, and of the misplacement of context.

The phenomena of localization have a bearing on the study of RI effects. Frequently, investigations of RI require recall of OL alone. Under these conditions, the data of recall contain materials from IL. The latter are called errors of intrusion, and have been interpreted as evidence of competition of responses. This interpretation rests, as we saw, on an inadequate analysis. The procedure in question places two requirements upon the *S*—to recall the items requested, and to differentiate between the lists—without clearly distinguishing between them. It ignores the possibility that *S* may produce "wrong" items not because of competition but rather because of uncertainty about list membership. The consequence may be to increase spuriously the reported levels of RI.<sup>13</sup>

The study of Postman (16), to which we referred in part earlier, further illustrates the importance of this distinction. It aimed in the main to trace the effects of interpolation on recognition of nonsense syllables. There were two recognition groups, each of which was first tested immediately following exposure of OL. The test was multiple-choice, each item from OL being shown together with three new syllables. Following the exposure of IL there was a second multiple choice recognition test again to select the OL item. The second recognition test also consisted of sets of four syllables, of which one came from the OL list. For Group 1, two of the items were from the first test of recognition, and one syllable was entirely new; for

<sup>13</sup> This procedure may also contribute to RI effects in relearning. On the first relearning trial, *S* is asked to recall OL items. He may know them, but fail to recall to which list they belong.



Group 2, the latter syllable was replaced by one from IL. RI was significantly higher for Group 2 than for Group 1. Postman concluded that the decrement obtained in Group 1 was due mainly to unlearning, while that of Group 2 was a consequence also of response competition.

The procedure and interpretation fail to separate the recognition of an item and the recognition of its context. On the final test, the Ss in Group 2 had to choose one from among four items, each of which they had seen previously, and in three different contexts. Since their task was to select only the item from OL, they had in effect to discriminate between one context and two others. They did less well than Group 1, in part because an appreciable proportion of their errors consisted of choices of IL items. We conclude that Postman did not test for recognition proper; to do so would have required a measure of recognition independent of accuracy of localization. The Ss of Group 2 might have recognized all items in a given set of four, without being able to localize them properly. Under these circumstances one cannot accept the inferences about the role of unlearning and competition of response in recognition.

Finally, the function of localization has a fundamental bearing on the process of recall. The occurrence of free recall presupposes a reference to the context of the materials recalled, and the presence of a connection between context and the data of recall. The properties of this connection are at present unknown. It is customary to consider it an instance of association, and to speak of association with context; this formulation is rather an extrapolation of accepted ideas than a product of investigation.<sup>14</sup> Further investigations of localization are now in progress.

### 5. Principles of Availability

The conceptual and empirical distinction between association and availability opens up the possibility for direct investigation of the determinants of availability. Earlier we have described the relevant findings derived from this investigation. We will now consider the problem in the light of related evidence.

a. *The principle of frequency.* Frequency is, we saw, a fundamental con-

---

<sup>14</sup> The context remains essentially unspecified when one describes it as a stimulus with which contents are associated. Therefore, to speak of association with context is to do little more than to refer to the free recall process as such. Further, associations with context imply the formation of multiple connections to a given stimulus, which is a condition of high associative interference. The rapid learning obtained under the present conditions (Figures 1 and 2) provides little evidence for such interference.

dition of availability. Let us now examine the further conclusion that the effect of frequency cannot, at least under the present conditions, be derived from associative considerations.

Many psychologists have held the latter interpretation of frequency, although conclusive investigation has been lacking. The effect of sheer familiarity has been asserted by Robinson (18, p. 118) and also by Thorndike (21, p. 63, 351). Recently, however, Deese (7) has proposed a purely associative interpretation of the effect of frequency on free recall that contradicts the conclusion here reached. His principal finding is that word lists of higher frequency yield higher recall. Deese questions whether "... there is any essential relationship between word frequency and free recall" (7, p. 23). Instead he proposes to construe the obtained differences in free recall in terms of an "interitem associative index," the argument being that lists composed of items which are (free) associates of each other will tend to yield higher recall scores the greater the number of interitem associations: "... the tendency for items to elicit one another as free associates is accompanied by a proportionate increase in free-recall scores" (7, p. 17).

The interpretation of Deese meets a difficulty when applied to the present findings; since associations between syllables were excluded, these cannot account for the effects of frequency. It is also conceivable that inter-item associations in no way influenced the recall of words. We may understand Deese's results by noting that an item belonging to a high associative index list was more likely to be recalled at the time of test for the same reason that it was more likely to be given as a free associate in the first place. Both effects are, we believe, directly traceable to the same underlying factor of availability. The attempt to comprehend the nature of free recall in terms of specific associations is, from the present point of view, to miss the singular importance of free recall as a unique process, one deserving of study in its own right.

The effects of frequency are a function of the specific conditions of acquisition. Earlier findings suggest that the more active the procedure of acquisition the higher is the subsequent recall (8, 23). Further evidence comes from the direct measure of availability by means of free recall. A recent investigation of Asch and Ebenholtz (1) demonstrated that free recall of paired associates following practice to a criterion is significantly higher for "response" terms, which were anticipated during learning, than for "stimulus" terms, which were only recited. Further, it was shown that recall of recited terms also increased with practice.



*b. Recency and primacy.* The effects of serial position that we have reported confirm previous findings (3, 25), and particularly those of Deese and Kaufman (5) and Deese (6). Our contribution has been to demonstrate that these serial effects are not a consequence of associative operations. There is reason to hold that recency is a condition that heightens availability, that it accounts for rate and order of acquisition, and also for relative ease of relearning. The effect of primacy may have to be considered separately; one must leave open whether primacy and recency in free recall refer to the same process.

*c. Isolation.* From the standpoint of the present analysis, isolation, such as von Restorff studied (24), is a condition of availability. Primacy may, under the present conditions, be an instance of isolation.

*d. Interpolation.* The effects of interpolation have been considered earlier.

The present account was not intended to be exhaustive, but rather to illustrate that the distinction between availability and association permits a more direct analysis of the conditions of availability.

#### E. SUMMARY AND CONCLUSIONS

Central to an associative theory of recall is the proposition that associative connections between the data of past experience alone determine the course of recall. This proposition implies that acquisition, too, of necessity consists of the formation of consecutive associations.

The object of this investigation was to explore the properties of acquisition and recall under conditions that excluded or reduced the formation of specific inter-item associations. The task consisted of the mastery of lists of nonsense syllables, the order of which was varied on successive trials. Acquisition and retention were measured by free recall.

Under these conditions acquisition (as measured by immediate free recall) was not substantially a function of the order in which items appeared during the learning trials. At the same time acquisition occurred with considerable rapidity. Nor was subsequent retention, which was also high, significantly a function of the order of initial experience. We concluded that short-term and long-term free recall occurred in the absence of specific inter-item associations.

The following determinants of differential recall were identified. (a) Serial primacy and recency. Items from the ends of the learning series occurred more frequently in immediate free recall than items from other positions. They were also recalled earlier than other items. Serial primacy and recency were thus conditions of recall and of order of free recall. (b) Long-term

recall was a function of frequency of recall during acquisition. (c) Employing an RI design, it was shown that interpolation produced substantial recall decrements. Interpolation was thus a negative condition of recall.

None of the preceding effects can be accounted for on the basis of specific inter-item associations. We proposed an interpretation of these effects in terms of availability or of accessibility of an item to recall. Availability refers to a condition of the memory trace, which must be distinguished from the concept of association, with which it has been confused. On this basis we drew the following conclusions: (a) Serial primacy and recency are conditions of availability. Frequency and order of recall are indices of availability, not of association or of habit strength. (b) RI obtained under the free recall conditions is a function of decreased availability due to interpolation. It cannot be interpreted in terms of current theories of interference, which are purely associative. Neither the competition of response hypothesis nor that of unlearning can account for the decrements obtained under the present conditions. Since impaired availability due to interpolation must also play a part in RI investigations employing the associative design, the latter findings must be reexamined from the standpoint of the concept of availability. The question was raised whether RI phenomena are under all conditions effects on availability, rather than on the associative bond. Although the general validity of this hypothesis is not established, it requires a reexamination of interference theories that were designed to fit an exclusively associationistic model of acquisition and recall.

The conceptual and empirical differentiation between availability and association permits a systematic analysis of the principles of availability. Availability is best studied by the procedures of free recall, uncomplicated by the essentially irrelevant introduction of associations. Such investigation should also help to clarify the relations between the processes of availability and association.

The procedure of interpolation provided preliminary information concerning the localization of recalled data in their appropriate contexts. Recalled items were predominantly accurately localized, although there were instances of incorrect reference. The process of localization may have a fundamental relevance to recall in general; the nature of the process awaits investigation.

There are properties of the recall process that cannot be explored except by means of free recall. Further investigations of the properties of free recall are now in progress.



## REFERENCES

1. ASCH, S. E., & EBENHOLTZ, S. M. The principle of associative symmetry. *Proceed. Amer. Philos. Soc.*, 1962, **106**, no. 2.
2. BARNES, J. M., & UNDERWOOD, B. T. Fate of first-list associations in transfer theory. *J. Exp. Psychol.*, 1959, **58**, 97-105.
3. BOUSFIELD, W. A., COHEN, B. H., & SILVA, J. G. The extension of Marbe's law to the recall of stimulus-words. *Amer. J. Psychol.*, 1956, **69**, 429-433.
4. CERASO, J. An experimental critique of competition of response and specific interference as factors in retroactive inhibition. Unpublished doctoral dissertation, New School for Social Research, 1959.
5. DEESE, J., & KAUFMAN, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. *J. Exp. Psychol.*, 1957, **54**, 180-187.
6. DEESE, J. Serial organization in the recall of disconnected items. *Psychol. Rep.*, 1957, **3**, 577-582.
7. ———. From the isolated verbal unit to connected discourse. In C. N. Cofer (Ed.), *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill, 1961, (Ch. 2).
8. GATES, A. I. Recitation as a factor in memorizing. *Arch. Psychol.*, 1917, **6**, No. 40.
9. GLAZE, J. A. The association value of non-sense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-269.
10. HULL, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730-734.
11. KRUEGER, W. C. F. The relative difficulty of nonsense syllables. *J. Exp. Psychol.*, 1934, **17**, 145-153.
12. MANDLER, G. Associative frequency and associative prepotency as measures of response to nonsense syllables. *Amer. J. Psychol.*, 1956, **68**, 662-665.
13. MCGEOCH, J. A. *The Psychology of Human Learning*. New York: Longmans, Green, 1942.
14. MELTON, A. W., & IRWIN, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173-203.
15. NOBLE, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421-430.
16. POSTMAN, L. Retroactive inhibition in recall and recognition. *J. Exp. Psychol.*, 1952, **44**, 165-169.
17. POSTMAN, L., & RAU, L. Retention as a function of the method of measurement. *U. Calif. Publ. in Psychol.*, 1957, **8**, 217-270.
18. ROBINSON, E. S. *Association Theory Today*. New York: Century, 1932.
19. RUNQUIST, W. N. Retention of verbal associates as a function of strength. *J. Exp. Psychol.*, 1957, **54**, 269-275.
20. SLAMECKA, N. J., & CERASO, J. Retroactive and proactive inhibition of verbal learning. *Psychol. Bull.*, 1960, **57**, 449-475.
21. THORNDIKE, E. L. *The Fundamentals of Learning*. New York: Teachers College, Columbia Univ., 1932.
22. UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59** (3, Whole No. 273).
23. UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and Verbal Learning*. New York: Lippincott, 1960.
24. VON RESTORFF, H. I. Über die Wirkung von Bereichsbildung im Spurenfeld (In W. Köhler & H. von Restorff, *Analyse von Vorgängen im Spurenfeld*), *Psychol. Forsch.*, 1933, **18**, 299-342.

25. WELCH, G. B., & BURNETT, C. T. Is primacy a factor in association formation?  
*Amer. J. Psychol.*, 1924, **35**, 396-401.

*Department of Psychology  
& Education  
Swarthmore College  
Swarthmore, Pennsylvania*

*Department of Psychology  
Connecticut College  
Storrs, Connecticut*



## TIME JUDGMENT AND SCHIZOPHRENIA: STEP INTERVAL AS A RELEVANT CONTEXTUAL FACTOR\*<sup>1</sup>

*Houston State Psychiatric Institute and Baylor University College of Medicine*

DAVID J. WRIGHT, SANFORD GOLDSTONE, AND WILLIAM K. BOARDMAN

### A. PURPOSE

Several writers (3, 5) have reported that schizophrenic Ss overestimate auditory durations when comparing the input with an internal standard, one clock second. When ascending and descending series of arithmetically spaced durations were judged as either *more* or *less* than one second,\* it was found that patients identified a shorter duration of input as equivalent to their concept than did healthy controls. These findings were interpreted as an experimental analogue of the disordered time judgment observed clinically within schizophrenic populations (2) reflecting a general disturbance in time judgment independent of the experimental context (3).

The present study altered one property of the experimental context, the step interval, hypothesizing that the previously used arithmetic spacing contained intrinsic anchors that tended to increase stimulus overestimation regardless of the dimension of judgment. Examination of the asymmetrical arithmetic step interval reveals an unequal psychophysical distance between adjacent stimuli involving greater subjective spacing at the lower end of the physical scale. It is possible that the asymmetry of the arithmetic series provided additional short anchors (1), resulting in the overestimation of a standard unit of time. This hypothesis is consistent with the increased anchor vulnerability of schizophrenic Ss using other dimensions (4) and the increased effectiveness of immediate anchors with time judgment (5).

This investigation compared the time judgments of schizophrenic and healthy Ss using both arithmetic and geometric stimulus intervals, the latter presumably involving less psychophysical asymmetry. The judgments of actively psychotic and more intact patients were compared by including a group of schizophrenics in remission.

\* Received in the Editorial Office on February 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Supported by U.S.P.H.S. Grant M-1121 and the Hogg Foundation.

## B. METHOD

### 1. Procedure

An electronic timer activated auditory durations with a reliable and accurate range of .01\* to 9.99 seconds. An audio-oscillator and headphones provided a tone of 725 cps and comfortable amplitude. Ss were presented a series of sounds using a modified method of limits procedure with either an arithmetic step interval of .1 second, or a geometric step interval of 10 per cent. S responded to each duration with the report *more* or *less* indicating a judgment of more or less than 1 second. Each arithmetic and geometric test consisted of ten alternating ascending and descending runs starting at 1.0 second. If S reported *more* to the initial 1.0 second stimulus, the first run was of a descending nature, whereas if S reported *less* to the initial stimulus, the first run ascended. Each ascending run terminated upon three consecutive reports of *more* and each descending run terminated upon three consecutive reports of *less*. The ascending and descending runs were continuous.

The percentages of each duration judged *less* for the ten runs were plotted on arithmetic probability paper and a straight line visually fitted. The Second Estimation Point (SEP) or duration which S reported *more* and *less* than one clock second 50 per cent of the time was derived from these plots.

### 2. Subjects

Fifty-four schizophrenic Ss between 14-50 years of age, of average intelligence or better, and without brain pathology or sensory deficit were studied. Thirty-six Ss were actively psychotic at the time of testing as defined by psychiatric appraisal of currently existing symptoms. Eighteen Ss were considered in remission as defined psychiatrically in terms of an absence of symptoms and the awaiting of discharge. Forty healthy control Ss were obtained from a middle class church and apartment development.

All Ss received both the arithmetic and geometric series alternately with order counterbalanced.

### 3. Groups

The patient and control Ss were divided into the following groups: Arithmetic followed by geometric series: Schizophrenia-Active,  $N = 19$ ; Schizophrenia-Remission,  $N = 10$ ; Healthy,  $N = 20$ . Geometric followed by arithmetic: Schizophrenia-Active,  $N = 17$ ; Schizophrenia-Remission,  $N = 8$ ; Healthy,  $N = 20$ .



## C. RESULTS

The Mann-Whitney and Wilcoxon tests were used to examine the differences between SEP distributions. Since no significant effect of presentation order was obtained, all arithmetic and all geometric data for a given population were combined. Table 1 summarizes the Median and Q values for the arithmetic and geometric SEP measures for each population. Since the non-parametric tests used in this report compare distributions, the

TABLE 1  
MEDIAN SEP AND Q VALUES OF THE HEALTHY, SCHIZOPHRENIA-ACTIVE,  
AND SCHIZOPHRENIA-REMISSION GROUPS FOR THE ARITHMETIC  
AND GEOMETRIC SERIES

	N	Arithmetic		Geometric	
		Mdn.	Q	Mdn.	Q
Healthy	40	.60	.22	.57	.20
Schiz.-Active	36	.43	.23	.60	.23
Schiz.-Remission	18	1.08	.46	.88	.37

Medians are useful only as points of reference. Table 2 summarizes the results of relevant Mann-Whitney and Wilcoxon analyses.

1. *Schizophrenia-Active and Healthy*

A significant difference between arithmetic and geometric series SEP measures was obtained for the Schizophrenia-Active ( $P < .001$ ) and Healthy ( $P < .05$ ) groups indicating that step interval was a relevant factor in these time judgments. The Schizophrenia-Active group obtained significantly ( $P < .05$ ) shorter SEP values with the arithmetic series than did the Healthy Ss, while no patient-control difference was obtained with the geometric series. The prediction that a more psychophysically symmetrical series would attenuate the previously obtained (3, 5) schizophrenia-Control SEP difference was verified, and this study replicated the shorter SEP for schizophrenic Ss with an arithmetic series.

2. *Schizophrenia-Remission*

This group showed unexplainably longer SEP values under both step interval conditions. Schizophrenia-Remission Ss differed from Healthy ( $P < .01$ ) and Schizophrenia-Active ( $P < .001$ ) Ss with the arithmetic series, and also from Healthy ( $P < .01$ ) and Schizophrenia-Active ( $P < .01$ ) with the geometric series. No arithmetic-geometric difference appeared.

TABLE 2  
CONFIDENCE LEVELS, MANN-WHITNEY AND WILCOXON TESTS

	Healthy Arithmetic	Healthy Geometric	Schiz.- Active Arithmetic	Schiz.- Active Geometric	Schiz.- Remission Arithmetic	Schiz.- Remission Geometric
Healthy Arithmetic	—	.05	.05		.01	
Healthy Geometric	.05	—		X*		.01
Schiz.-Active Arithmetic	.05		—	.001	.001	
Schiz.-Active Geometric		X*	.001	—		.01
Schiz.-Remission Arithmetic	.01		.001		—	X*
Schiz.-Remission Geometric		.01		.01	X*	—

\* Not Significant



#### D. DISCUSSION

These findings emphasize the importance of contextual factors in determining judgmental differences between patient and healthy populations. The significantly shorter SEP obtained earlier (3, 5) can be accounted for in terms of a differential effect of step-interval-determined anchors. Healthy Ss were influenced by these intra-serial anchors but not as much as the actively schizophrenic patients. Excessive vulnerability to anchors represents an experimental-psychological analogue of the passivity of the schizophrenic patient who is dominated by contemporary external inputs. It is likely that he is less able to mobilize frames of reference or concepts from his past to do battle with immediate background forces.

No simple explanation is available for the longer SEP values of the schizophrenic Ss who were in remission. The possibility that they are compensating for excessive vulnerability to anchors, or may possess an "input hunger" should be studied further.

This report demonstrates the danger of generalizing a finding obtained within a highly specific experimental context to an entire dimension of judgment.

#### E. SUMMARY

The previous finding of greater overestimation of short auditory durations by schizophrenic Ss was reexamined by studying the differential effect of an arithmetic and a geometric stimulus series on Healthy ( $N=40$ ), Schizophrenia-Active ( $N=36$ ), and Schizophrenia-Remission ( $N=18$ ) Ss. The following results were obtained:

1. The Healthy and Schizophrenia-Active Ss demonstrated a difference between the two series. Judgments with the arithmetic series acted as if they were influenced by short anchors as predicted.
2. The previously found difference between Healthy and Schizophrenia-Active Ss was obtained with the arithmetic interval. The patients behaved as if they were more influenced by short anchors.
3. No difference between Healthy and Schizophrenia-Active Ss was found with the geometric series.
4. The Schizophrenia-Remission Ss mysteriously obtained longer judgments than the other two groups.

These results are discussed in terms of differential anchor vulnerability and the need to consider specific contextual factors as the crucial variables in the study of psychopathology.

## REFERENCES

1. GUILFORD, J. P. *Psychometric Methods*. (2nd ed.) New York: McGraw-Hill, 1954.
2. LEWIS, A. Experience of time in mental disorder. *Proc. Roy. Soc. Med.*, 1932, 25, 611-620.
3. LHAMON, W. T., & GOLDSTONE, S. The time sense: Estimation of one second durations by schizophrenic patients. *Arch. Neurol. & Psychiat.*, 1956, 76, 625-629.
4. SALZINGER, K. Shift in judgment of weights as a function of anchoring stimuli and instructions in early schizophrenics and normals. *J. Abnorm. & Soc. Psychol.*, 1957, 55, 43-49.
5. WEINSTEIN, A. D., GOLDSTONE, S., & BOARDMAN, W. K. The effect of recent and remote frames of reference on temporal judgments of schizophrenic patients. *J. Abnorm. & Soc. Psychol.*, 1958, 57, 241-244.

*Department of Psychiatry*  
*Baylor University College of Medicine*  
*Houston, Texas*



## FREUD'S CONCEPTION OF THE ORIGIN OF CURIOSITY\*

*Department of Psychology, Brandeis University*

---

JOEL ARONOFF

---

### A. GENERAL PURPOSE

Psychological theorists have come, in the past few years, to an examination of exploratory behavior in animals and curiosity in men in hopes of discovering in this activity a means of advancing and broadening their theoretical understanding. Reports of this behavior, which had previously been dismissed as sentimental or anecdotal, have now come to be seen as crucial in developing psychological theory so that it might provide a means of including all the many complexities of normal human existence. In this effort a great many review articles and exhortatory statements have been published, all pointing out the failure of drive-reduction theory to explain human behavior, and stating the great advantages to be gained by re-interpretation of the very acts which have driven these theorists, in the first place, to seek and to inquire.

There is always place set aside in these reviews for at least a mention of Freud's ideas on this topic. Unfortunately, this often turns out to be only a bow to the east and a citing of scripture, without a full understanding of what Freud had to say and why he was brought to such conclusions. For this reason, I feel that at a time when a reordering of psychological knowledge is occurring and new concepts are being built upon the old, it is very important to comprehend fully the contribution Freud made to this problem and the context in which it must be understood.

At the outset it is most important to realize that Freud never addressed himself directly to the problem at all. The only way to discover exactly what he felt, is to go through his writings and draw together a general and comprehensive statement from his many fragmentary allusions. It is not surprising that his ideas are to be found in that state, because his theory was developed in order to explain the immediate clinical problems, and not the more general one of human curiosity.

Because Freud never specifically attempted to organize his thoughts on this particular question, what emerges from his writings is an interesting

---

\* Received in the Editorial Office on February 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

confrontation of two directly opposed theoretical statements; which reflect in essence a certain dualism to be found in much of his work. This is the universal problem of the biological and the social. While his whole philosophy and explication of human activity is based on the vicissitudes of an innate drive system, his understanding of the cognitive apparatus of man and its functions was specifically stated to be a societal creation; a system created and developed solely by the challenges which society presents to the organism. For this reason, close study of his work shows that at certain times Freud considered the curiosity of man to be due to a biological urge, and at other times to be a coping mechanism created to master the problems of life.

I would like in this paper to describe each of Freud's ideas on the origin of human curiosity, and by presenting a number of quotations taken from his works illustrate clearly the distinct nature of each concept and the extent to which it contrasts with the other.

## B. THEORIES OF CURIOSITY

### 1. *Biological Drives*

The first of Freud's ideas on the origin of human curiosity grew out of his general conception of drives. Throughout his work the main emphasis rested always on somatically based drives which energize behavior. As the organism passes through the various pregenital stages, successive patterns of activity emerge as the result of maturation of the organs which are the sources of the libido. In the anal-sadistic stage of development the energy searching for release has, as one of its component aims, an instinct for looking which Freud called scopophilia. This partial instinct is a natural occurrence of man's biological organization when it reaches this level of development. Scopophilia was seen by Freud to be a drive which causes the child to develop an intense interest in sexual matters. The child is impelled to look, search for and explore the sexual organs of his parents, playmates, and even the inanimate objects in his environment. A good example of this would be to recall the way Freud portrayed little Hans ("A Phobia in a Five Year Old Boy"), who was fascinated by his penis (his "widdler"), and whose immediate impulse when meeting any person or object was to examine it to see if it too had a widdler.

Sexual curiosity starts, according to this first formulation, by the child discovering and taking pleasure in the sensations of his sexual organ. This source of pleasure becomes associated with the partial instinct of scopophilia which develops at the age of three to five. Through this association Freud extended his biological theory of human activity to include an explanation



for what he saw as the child's intense interest in sex. It is also true that Freud often used this theoretical exposition as a means of illustrating his point that children are sexual creatures (2).

Freud felt that through this activity children come to investigate the world and are able thus to arrive at much abstract knowledge. For example, little Hans was able to arrive at the distinction between animate and inanimate objects by exploring them to see if they possess a "widdler." Since Freud was not concerned with human curiosity, *per se*, but rather with explaining the set of phenomena with which he was involved (children's sexual questioning), he did not extend his discussion to a detailed analysis of adult curiosity. The answer he gave when faced with this question was to say simply that a pattern of enquiry had been set up and can be found as a derivative instinct in adulthood. The only other time Freud dealt with adult curiosity was when he demonstrated that the obsession for understanding was due to repression of the looking impulse (4).

Several direct references to Freud's writings will clearly reveal the biological aspect to be found in his work on curiosity.

Thirst for knowledge seems to be inseparable from sexual curiosity. Han's curiosity was particularly directed towards his parents. (5, p. 153)

The pleasure which a person takes in his own sexual organ may become associated with scopophilia . . . in its active and passive forms. . . . So little Hans began to try to get a sight of other people's widdlers; his sexual curiosity developed, and at the same time he liked to exhibit his own widdler. (5, p. 248)

At the beginning of its activity the scopophilic instinct is auto-erotic: . . . that object is the subject's own body. It is only later that the instinct comes . . . to exchange this object for the analogous one of the body of another. (6, p. 72)

Impulses with a passive aim are connected with the eroto-genic zone of the rectal orifice, at this period very important; the impulses of scopophilia (gazing) and curiosity are powerfully active. . . . (8, p. 287)

Partial Impulses . . . the infantile sexual life . . . shows components for looking, showing off, and for cruelty, which manifest themselves somewhat independently of the erotogenous zones. . . . The little child . . . evinces definite pleasure in displaying his body and especially his sexual organs. A counterpart to this perverse desire, the curiosity to see other persons' genitals, probably appears first in the latter years of childhood when the hindrance of the feeling of shame has already reached a certain development. Under the influence of seduction, the looking perversion may attain great importance for the sexual life of the child. . . . Still, from my investigations of the childhood years of normal and neurotic patients, I must conclude that the impulse for looking can appear in the child as a spontaneous sexual manifestation. (1, p. 593)

The boy is excited and stimulated by his genitals—assumes all living things are the same as he

. . . indeed we know that he investigates inanimate objects with a view to discovering something like his member in them. . . . He wants to see the same thing in other people, so as to compare it to his own. . . . The driving force which this male portion of his body will generate later at puberty expresses itself in childhood essentially as an impulse to inquire into things—as sexual curiosity. (9, p. 246)

Inquisitiveness . . . from the age of three to five, there appear the beginnings of that activity which are ascribed to the impulse for knowledge and investigation. The desire for knowledge can neither be reckoned among the elementary instinctive components, nor can it be altogether subsumed under sexuality. Its activity corresponds, on the one hand, to a sublimated form of acquisition, and on the other hand, the energy with which it works comes from the looking impulse. . . . The inquisitiveness of children is directed to sexual problems unusually early . . . indeed, curiosity may perhaps first be awakened by sexual problems. (1, p. 594)

The child was under the domination of a component of the sexual instinct, scopophilia (the instinct of looking), as a result of which there was a constant recurrence in him of a very intense wish connected with persons of the female sex who pleased him—the wish, that is, to see them naked. (4, p. 300)

There can be no doubt as to the existence of Han's sexual curiosity; but it roused the spirit of inquiry in him and enabled him to arrive at genuine abstract knowledge. (5, p. 153)

This interest (in his widdler) aroused in him the spirit of inquiry, and he thus discovered that the presence or absence of a widdler made it possible to differentiate between animate and inanimate objects. (5, p. 248)

## 2. *Social Coping*

A totally different explanation of the origins of human curiosity emerges as well from an examination of Freud's writings. While the vicissitudes of the instincts are used to explain most of human motivation, the cognitive structure of man is not seen as naturally evolving through a process of growth. It is rather an apparatus forced upon man to enable him to cope with the rigors of living in society. These somewhat alien processes are developed primarily as a problem solving mechanism, which is tolerated for the pleasures it can ultimately achieve for the individual. In his essay on "Wit" Freud shows quite clearly how he believed thought to develop. The child takes great pleasure in saying words—solely to enjoy their rhyme and rhythm. Gradually society deprives the child of this pleasure until only the useful and senseful connection of words is allowed him. These "mental inhi-



bitions" are very distasteful to him, and he finds relief in splees, songs and humor at every opportunity. After a scientist makes a speech at a scientific convention he searches for relaxation from the strain by reading the funny papers.

It was within this framework of a system of cognitive processes, which had been forced upon the individual, that Freud developed his second theory of human curiosity. The problem which had attracted Freud was children's incessant theorizing about the origins and mechanisms of birth. Freud felt that the answer to this lay in the jealousy which each child must feel at the birth of a rival. This newcomer presents a great threat to the child's privileged position in the family, for it necessitates his sharing the affection and love of the parents. In order to overcome this danger, and to prevent the occurrence of another such calamity, the child begins to search and to enquire into the causes of this catastrophe. It is this specific problem which develops in the child the capacity of curiosity. If his questions had been answered early this inquisitiveness would never have developed. The threat would have been removed, and the child would have gone on to occupy himself with other things.

This conception, which is in direct contrast to the former, explains curiosity as being basically a coping mechanism. In this formulation there is no question of curiosity being an instinctive aspect of the personality. Here it is merely the reaction of an individual to a threat.

Once again, we can clearly see this concept develop by turning to Freud's writing and examining a number of passages which are directly concerned with this topic.

It is not theoretical but practical interests, which start the work of the child's investigation activity. The menace to the conditions of his existence through the actual or expected arrival of a new child, the fear of losing the care and love connected with this event, cause the child to become thoughtful and sagacious. . . . the first problem with which he occupies himself is not the question of the difference between sexes, but the riddle: Where do children come from? (1, p. 595)

Children do not start investigating the difference between sexes—they know it exists.

The child's desire for knowledge does not awaken spontaneously at this point at all, as it would if prompted perhaps by an inborn need to seek for causes, but arises under the goad of a selfseeking impulse which dominates him when he is confronted by the arrival of a new child—perchance at the end of the second year. Those children whose own nursery at home does not become divided up in this way are

nevertheless able as the result of their own observations to put themselves in the place of others who are in this situation in other homes. The loss of the parents' care and concern . . . has the effect of awakening the emotions of the child and sharpening its thinking capacities. (3, p. 62)

Under the stimulus of these feelings and anxieties the child thus comes to consider the first of the great problems of life, and asks itself the question where children come from. . . . The question itself, like all inquiry, is a product of dire necessity. (3, p. 62)

The second great problem which exercises a child's mind—probably at an earlier date—is that of the origin of children, and is usually aroused by the unwelcome arrival of a baby brother or sister. This is the oldest and most burning question that assails immature humanity . . . . The answers given to children in the nursery wound the child's frank and genuine spirit of investigation. (2, p. 40)

I must still say something about sexual curiosity in children. It is too characteristic of childish sexuality and too important for symptom-formation of the neuroses to be omitted. Infantile sexual curiosity begins very early, sometimes before the third year. It is not connected with the difference between the sexes, which is nothing to children, since they—boys at least—ascribe the same male genital organ to both sexes . . . .

The sexual interest of children is primarily directed to the problem of birth—the same problem that lies behind the riddle of the Theban Sphinx. This curiosity is for the most part aroused by egoistic dread of the arrival of another child. The answer which the nursery has ready for the child, that the stork brings the babies, meets with incredulity even in little children much more often than we imagine. The feeling of having been deceived by grown-up people, and put off with lies contributes greatly to a sense of isolation and to the development of independence. (8, p. 278)

Hans keeps cudgelling his brains to discover what a father has to do with his child, since it is the mother who brings it into the world. (5, p. 242)

We find ourselves obliged to take as our starting point an outburst of sexual pleasure and sexual curiosity connected like this one, with the birth of the next child. (5, p. 255)

It is much more important that children should never get the idea that one wants to make more of a secret of the facts of sexual life than of any other matter not suited to their understanding. To insure this it is necessary that from the very beginning everything sexual should be treated like everything else . . . . In this way the curiosity of children will never become very intense, for at each stage in its inquiries it will find the satisfaction it needs. (2, p. 42)

### C. CONCLUSION

Discovering two contrasting explanations for the same phenomena should not appear surprising, nor should it be disconcerting, for each concept was



developed in response to a particular line of reasoning, and the general problem of human curiosity was never specifically explored for its own sake at all. Each of his concepts had equal importance to Freud in the context in which it was developed. What is important is that when we turn to Freud for guidance for our own thinking, we understand clearly all that he had to say on curiosity, and only then will we be able to evaluate and integrate it into our own theorizing.

## REFERENCES

1. FREUD, S. Three Contributions to the Theory of Sex. Modern Library, 1905.
2. ———. The sexual enlightenment of children. *Collected Papers*, 2, 1907.
3. ———. On the sexual theories of children. *C.P.*, 2, 1908.
4. ———. Notes upon a case of obsessional neurosis. *C.P.*, 3, 1909a.
5. ———. Analysis of a phobia in a five-year-old boy. *C.P.*, 3, 1909b.
6. ———. Instincts and their vicissitudes. *C.P.*, 4, 1915.
7. ———. From the history of an infantile neurosis. *C.P.*, 3, 1918.
8. ———. A General Introduction to Psycho-Analysis. Garden City, 1920.
9. ———. The infantile genital organization of the libido. *C.P.*, 2, 1923.

*Department of Psychology*  
*Brandeis University*  
*Waltham, Massachusetts*

# ATTITUDES OF A GROUP OF NURSING STUDENTS TOWARD A CLINICAL AFFILIATION IN A PUBLIC HOSPITAL FOR THE MENTALLY ILL\*<sup>1</sup>

*University of Buffalo*

---

MILDRED E. FLETCHER

---

## A. INTRODUCTION

Today, over ninety per cent of the schools of nursing in the United States provide psychiatric nursing experience in the basic schools of nursing (1, pp. 81-91). With the advancement of modern psychiatric practice, the importance of positive attitudes toward the nursing care of psychiatric patients is recognized by the nursing profession. Educators in the field of psychiatric nursing, particularly, are concerned with the development of such attitudes by students and their ability to adjust to psychiatric patients in a therapeutically useful manner (1, p. 4). Weiss affirms that "Nursing care involves the use of attitudes in giving patients the greatest opportunity to regain their health and to learn good health habits (21, p. 3)."

The increasing emphasis on attitude therapy makes it apparent that the psychiatric nursing program must take into account the importance of the attitudes of the student in establishing a satisfactory nursing experience. The study of Black (1, pp. 4-17) showed that many groups who are concerned with nursing education are becoming increasingly aware that the human relations content of the basic curriculum can become a vitalizing part of the clinical areas and that there is need for further research in teaching nursing students new interpersonal skills and attitudes.

Some studies point up the significant relationship of the attitude of the student and the outcome of the educational experience. Robinson (14, pp. 1091-92) and Cavaglieri (2, pp. 212-13) identified the phenomena of resistance to psychiatric nursing in the attitudes of student nurses which appear to be reflected in retarding their adjustment to the psychiatric nursing experience. Matheny (10, p. 686) identified the phenomenon of resistance

---

\* Received in the Editorial Office on March 12, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This article is a condensed report of a thesis which was carried out by the author in partial fulfillment of the requirements for the degree of Master of Science in Nursing at University of Buffalo, Buffalo, New York, 1958. The complete study is on file in the library of the University of Buffalo.



to psychiatric nursing and indicated that most of the students did not learn to understand normal behavior during the twelve-week affiliation.

Other studies explored the resistance to psychiatric nursing in the attitudes of students in anticipation of and before they begin this experience, feelings which appear to hinder their adjustment to psychiatric nursing. Coggins (3, pp. 1-34) and Wright (22, pp. 1-99) indicated that student nurses appear to have feelings of apprehension, fear and anxiety about mental illness prior to having experience in this field, and show the need for alleviation of such problems. Another study offered some clues on the reduction of negative feelings in students toward psychiatric nursing during the first five weeks of the affiliation. Samhammer (15, pp. 1-44) found that anxiety producing factors which appear to be present in the attitudes of students at the beginning of the psychiatric affiliation, tend to be reduced as students have increasing contacts with psychiatric nursing, and that there is need for understanding of these emotions by nursing educators.

Two studies further explored the attitudes of students by testing groups of them at the beginning and at the end of the affiliation. Galloway (4, pp. 1-68) indicated that the psychiatric experience had little effect in changing certain negative attitudes toward mental illness. Johnson (7, pp. 1-69) concluded that certain misconceptions and undesirable attitudes toward psychiatric nursing were increased, that some were changed to more favorable attitudes at the end of service, that negative attitudes should be recognized by nursing educators and that efforts should be made to change them constructively.

These studies just reviewed emphasized the importance of the initial attitudes of the student toward psychiatric nursing, and indicated that the emotional component of her attitude appears to affect her adjustment in the experience. These observations were made under various conditions. In this study, the emphasis is on learning what attitudes the affiliated students have toward psychiatric nursing. Although other areas may be observed, only those attitudes related to (a) preconceptions, (b) expectations, (c) personal relations, (d) personal feelings, and (e) attitudes and activities of patients with whom the student works, are to be studied.

## B. PURPOSE

During a two-year period, observation of approximately 400 student nurses from four different schools of nursing affiliated at the selected psychiatric hospital showed a certain pattern of resistance toward psychiatric nursing and/or toward the psychiatric affiliation. Socialized

discussion periods in the classroom gave opportunities to individual students to present problems. Many times their questions and statements as well as their attitudes, as indicated by facial expressions, posture, and tone of voice, reflected fear of and/or resistance to psychiatric nursing. The phenomena of resistance are familiar to instructors in the field (10, pp. 684-87). Help with the adjustment of the student to psychiatric nursing, per se, is dependent upon an understanding of her attitude. The problem, then, with which this study deals can be best summarized in terms of its purposes which were to determine: (a) the attitudes of nursing students toward psychiatric nursing; (b) the attitudes of student nurses toward the psychiatric affiliation; and (c) whether attitudes reveal implications relative to future program planning for students during the psychiatric nursing experience.

For this study, *attitudes* will be defined as "an emotional tendency, organized through experience, to react positively or negatively toward a psychological object (13, p. 363)." *Opinion* will mean verbal or written expression of an attitude (18, p. 7).

### C. BASIC ASSUMPTIONS

The study attempts to test the hypothesis that student nurses have attitudes toward psychiatric nursing and are able to reveal them in response to an attitude questionnaire. It is based upon the assumptions that:

1. Students have a knowledge of the preconceptions of society in regard to mental institutions, mental illness and mentally ill patients. The influence of these preconceptions may affect their opinions.

2. Students entering this new learning experience, have expectations concerning the content, methods of teaching, the hospital wards and the nursing activities. The anticipation of what others expect of them and their own aspirations may affect their opinions.

3. Students' feelings of insecurity during the dependency-independency stage of late adolescence may be increased by having to adjust personal relations with other people in an unfamiliar setting. This may affect their opinions.

4. Students may have feelings of guilt, and worry over their own personality effect on the mentally ill, and these feelings may be increased by their own estimate of success in adjustment to psychiatric nursing of the mentally ill. Such feelings may affect their opinions.

5. Student nurses entering the affiliation, may have negative feelings toward mentally ill patients, such as those of fear, rejection and frustration.



These feelings may be increased by the attitudes and activities of the mentally ill patients with whom they work. Thus, the attitudes and activities of these patients may affect the students' opinions.

#### D. SUBJECTS

The thirty-five student nurses who participated in the study, were regular students enrolled in the basic school of nursing programs of the three hospital schools of nursing affiliating with the psychiatric institution. The intelligence level of the group members was average. The test results on student nurses from one of the hospital schools were not available. The students ranged in age from 19 years and 3 months to 22 years and 2 months. They had been in nursing school between 18 and 31 months before entering upon their psychiatric affiliation. For some students the psychiatric engagement was their first experience away from their home school.

The students had no previous acquaintance with psychiatry other than the general psychology offered in the pre-clinical program. The majority of the students had never visited a mental hospital.

This brief description of the students in relation to maturity, age, intelligence, year in school, and contact with psychiatry shows the tentative characteristics of the subjects taking part in the study. It was assumed that the age of late adolescence with its stage of dependency, separation from the home school and the newness of the theory of psychiatry, all might have some effect on the attitudes and adjustment of the student to the psychiatric affiliation.

#### E. METHOD

A psychiatric nursing attitude questionnaire was constructed according to criteria set up by leading authorities in the field of attitude questionnaire construction, illustrated by Remmers and Gage (13, pp. 381-441), Thurstone and Chave (18, pp. 22-23), Thorpe, Clark and Tiegs (19), Koos (9, pp. 1-178), and Payne (11, pp. 1-249). Tentative areas of the objects of reference of attitudes toward psychiatric nursing and toward the psychiatric affiliation are illustrated in the work of Robinson (14, pp. 1091-92), Kalkman (8, pp. 7, 8; pp. 221-49), Coggins (3, pp. 1-34), Galloway (4, pp. 1-68), Wright (22, pp. 1-99), Johnson (7, pp. 1-69), Samhammer (15, pp. 1-44), Steele and Manfreda (17, pp. 20-21), Schwartz and Shockley (16, pp. 21-71), Zilboorg (23, pp. 3-371), Peplau (12, p. 326), and Wall (20, p. 401). The objectives of the attitude questionnaire were formulated from the data of these references.

The first form was given to eighteen student nurses during the latter part of their twelve-week psychiatric service to obtain suggestions as to the clarity of the items. A revised form was constructed by rewording some questions and by replacing some deleted items with more appropriate questions. The entire questionnaire used in the investigation consisted of 100 questions, divided into five areas relating to the attitude objects.

1. *Preconceptions*

- a. The mental institution
- b. Commitment of the patient
- c. Condition of mental illness
- d. Associates' opinions about mental illness
- e. The public's opinions about mental illness
- f. Association with the recovered mentally ill

2. *Expectations*

- a. Acquaintance with the practice field
- b. Status of psychiatric nursing
- c. Status of the mentally ill and his care
- d. Convulsive treatment
- e. Abnormality of psychotic behavior
- f. Acquaintance with the concepts of psychiatry

3. *Personal Relations*

- a. Interaction with environmental factors
- b. Dependency-independency
- c. Exercise of social standards
- d. Freedom from anti-social tendencies
- e. Relationship with family
- f. Exercise of social skills with normal people
- g. Nursing personnel relationship

4. *Personal Feelings*

- a. Reaction between self and the mentally ill
- b. Reaction between self and therapeutic treatment
- c. Reaction between self and psychiatric nursing
- d. Nervousness
- e. Personal attitude effect on the mentally ill
- f. Reaction to self-excusing

5. *Attitudes and Activities of Psychotic Patients*

- a. The freedom of the unaccompanied patient
- b. Frustration by psychotic behavior
- c. Fear of patient
- d. Reason for abnormal behavior
- e. Climate of student-patient communication

Each component was composed of twenty questions yielding evidences of



the presence or absence of an opinion of its kind. The hundred, (100) questions were designed to give positive, negative, and "I don't know" responses. The "Yes," "No," and "?" answer columns were considered adequate even though it was recognized that the "I don't know" was likely to be used as a blanket answer—an unwillingness to respond.

This revised questionnaire was taken by thirty-five student nurses from three hospital schools of nursing in Western New York, who had completed a twelve-week psychiatric affiliation in a large mental hospital during April 1, 1957 to June 22, 1957. It was administered in the following way:

1. Part I, Preconceptions, was given in individual conferences with each subject, during the first few days of their affiliation, and, again during the final week of affiliation. The responses to Part I were oral.
2. Part II, Expectations, Part III, Personal Relations, Part IV, Personal Feelings, and Part V, Attitudes and Activities of Patients, were given to all of the subjects in a group setting during the sixth week of the affiliation, and again, during the twelfth week at the termination of the affiliation. Responses to Parts II, III, IV, and V were written.

The results of the two administrations of all five parts of the attitude questionnaire were compared to ascertain what attitudes were present at the beginning, during the mid-term, and at the end of the affiliation.

Care was exercised in the administration of the questionnaire to secure thoughtful consideration of the test items and intellectually honest reactions to them. The students were told of the value of such schedules and their participation was encouraged. Also they were assured that the data were not related to school grades, that it was not a test which one could pass or fail, that their personal opinions were the best answers. Furthermore, a more favorable attitude toward the project was also sought by allowing the subjects to indicate the statements about which they were uncertain—the directions calling for a choice of one of three responses: "Yes," "No," and "?." All thirty-five students participated in the study.

## F. RESULTS

Each of the 100 statements of the questionnaire was considered to be either positive or negative in accordance with the writer's judgment of the probable interpretation given to it by the subjects and in accordance with the writer's conception of current psychiatric nursing knowledge or estimate of the consensus of authorities.<sup>2</sup> Responses that indicated freedom from unfounded beliefs, hopeful

<sup>2</sup> Personal interview with Eleanor A. Jacobs, Ph.D., Instructor in Psychology and Consultant in Clinical Psychology, University of Buffalo, Buffalo, New York.

expectations, healthy social relations, belief in self, and a wholesome respect for the mentally ill patients were considered favorable, positive ones. Responses that indicated biases, unhealthy outlook, unwholesome social behavior, insecurity, and a rejecting or authoritarian reaction to the mentally ill patients were considered unfavorable, negative ones. Responses of "I don't know" were scored negative responses because the direction of these responses indicated some reason for blocking, having difficulty in relating to the situation, or an inability to take a positive stand.<sup>3</sup>

The 100 questions checked by the 35 subjects were converted into numerical scores. The favorable response was assigned the positive value of +1, and the unfavorable response was assigned the negative value of -1. The scores were tabulated in the manner of a teacher-made test, and graded in accordance with a teacher-made key. The numerical values were designated in an arbitrary fashion and then verified objectively.

The score for the total group was determined by finding the average numerical values on the entire set of questions. The allotment of a definite value to each response made it possible to re-test the same group with the same set of questions to note any change or shifts in the expressed attitudes of the subjects.

The scores of the total group on the initial test ranged from 51 to 81, while on the post-test they ranged from 54 to 87. A Pearson correlation between the two sets of scores, ungrouped data in a double entry table, yielded an  $r$ , of .658, which is significant according to the standard error of  $r$ , when the population  $r$  is assumed to be zero. Since the  $r$  of .658 is 3.87 times this standard error, it can be assumed that the correlation is significant (6, pp. 180-81). Fisher's  $t$  Formula, used in the case of small samples, gave a  $t$  ratio of 4.87. Since this  $t$  ratio of  $r$  is 2.732 times the required  $t$ 's, the hypothesis that the population correlation is zero can be rejected, and  $r$  is significant (6, pp. 219-20, 539). Table 1 summarizes the measures of reliability just described.

In addition to the statistical analysis of the score results as interpreted in Table 1, the mean scores of grouped data on each of the five component parts were computed to determine changes or shifts in attitudes as expressed by differences between the first and final tests. The mean score, the standard deviation, and the differences in the means on each of the five parts of the questionnaire are illustrated in Table 2.

To determine whether the difference in the means of the two sets of scores

---

<sup>3</sup> *Ibid.*



TABLE 1  
COEFFICIENT OF CORRELATION; UNGROUPED DATA; AND OTHER STATISTICS OF  
TWO SETS OF SCORES OF THIRTY-FIVE NURSING STUDENTS ON A PSY-  
CHIATRIC NURSING ATTITUDE QUESTIONNAIRE ADMINISTERED  
TWO TIMES DURING AFFILIATION

Statistic	Initial test	Post-test
N	35	35
M	72	76
$\sigma$	-1.97	-2.42
$r$		.658
$\sigma r$		.17
$t$ of $r$		4.87

TABLE 2  
CHANGES IN AVERAGE SCORES ON FIVE PARTS OF ATTITUDE QUESTIONNAIRE:  
GROUPED DATA: INITIAL TEST AND POST-TEST, BY THIRTY-FIVE NURSING  
STUDENTS DURING TWELVE-WEEK PSYCHIATRIC AFFILIATION

Areas	Initial test		Post-test		Difference Means
	Means	SD	Means	SD	
Part I	20.5	1.62	23.6	.05	3.1
Part II	25.5	2.92	27.4	1.84	1.9
Part III	26.4	3.48	24.4	7.6	2.0
Part IV	27.0	2.6	27.2	7.34	.2
Part V	24.0	4.7	25.2	2.2	1.2
Sums	2,445		2,569		
Means	70.5		74.4		3.9
SD	4.4		3.9		

in Table 2 was significant the  $t$  test of significance between uncorrelated means was applied by use of Formula 29 and Formula 24A (5, p. 198). These results appear in Table 3. The differences between the two means in the areas of Personal Relations, Personal Feelings and Attitudes and Activities of Patients show the  $t$  value is at the one per cent level so that the hypothesis of chance can be rejected, for a real difference does exist. In the area of Pre-conceptions the difference is at the five per cent level, while for Expectations the  $t$  score difference is at the 10 per cent level. The  $t$  test of significance between the two mean scores of the two tests, as a whole, shows a significant difference in the attitude changes of the subjects during the twelve-week affiliation.

### G. DISCUSSION

This shift in attitude raises questions: (a) Were the attitudes changes increased favorable attitudes or decreased favorable attitudes? (b) What statements rated and/or effected the greatest attitude changes in the areas of Personal Relations, Personal Feelings, and Attitudes and Activities of Patients?

TABLE 3  
*t* TESTS OF SIGNIFICANCE FOR THE FIVE COMPONENT PARTS OF THE ATTITUDE  
 QUESTIONNAIRE, INITIAL TEST AND POST-TEST, BY THIRTY-FIVE NURSING  
 STUDENTS DURING PSYCHIATRIC AFFILIATION

Areas	Initial Mean	test SD	Post-test Mean	SD	Difference in Means	<i>t</i> score	Level of sig- nifi- cance
Part I (Preconceptions)	20.5	1.62	23.6	.05	3.1	1.93	.05
Part II (Expectations)	25.5	2.92	27.4	1.84	1.9	1.87	.10
Part III (Personal rela- tions)	26.4	3.48	24.4	7.6	2.0	4.18	.01
Part IV (Personal feel- ings)	27.0	2.6	27.2	7.34	.2	38.9	.01
Part V (Attitudes and activities of patients)	24.0	4.7	25.2	2.2	1.2	4.31	.01

In the area of Personal Relations, the smaller spread of ratings around the mean on the post-test signifies a decrease in the number of favorable responses. Since this component of the attitude questionnaire discloses whether or not the subjects' feelings of insecurity during the dependency-independency stage of late adolescence are increased by having to adjust personal relations with other people in an unfamiliar setting, it is possible that the ratings reflect these differences. The per cent of students showing the greatest change to unfavorable responses on the post-test are centered about the statements that (a) the subject does not find relief from anxiety by talking over problems with parents, (b) the subject does not feel at ease while working with the nursing personnel on the services.

In the area of Personal Feelings the *t* score of 38.9 shows a very real difference. With the mean score on the post-test only slightly higher than that on the initial test, the concentration of ratings around the mean signifies an increase in the number of favorable responses. Since this component discloses whether or not the subjects have feelings of guilt and worry over their own personality effects on the mentally ill and whether or not these feelings are increased by their own estimate of success in adjustment to psychiatric nursing, it is possible that concentrated mean ratings reflect these changes. The per cent of students which effected the greatest change to favorable responses on the post-test, are centered about the statements that: (a) the subject does



not worry over shock treatments the patients receive; (*b*) the subject is able to adjust to many changes in the nursing assignments on the clinical services; (*c*) the subject does not have trouble sleeping at night when the patients are making noise; and (*d*) the subject gains satisfaction in the care of the psychotic patient.

In the area of Attitudes and Activities of Patients, the *t* score between the initial and post-tests is significant. The higher mean on the post-test denotes a change by an increase in the number of favorable responses. Since this component discloses whether or not the subjects have negative attitudes toward the mentally ill patients, such as, feelings of fear, rejection, or frustration, and whether or not these feelings may be increased by the attitudes and activities of the mentally ill patients with whom the subjects work, it is possible that the ratings reflect these differences. The per cent of students which effected the greatest changes toward favorable responses on the post-test in this area, were centered about the statements that: (*a*) the subject would not cross the street when alone and seeing a patient on ground parole; (*b*) the subject is not afraid when alone and seeing a patient walking freely about the hospital grounds; (*c*) the subject would not agree with a patient that he is a sinful person, when he complains very bitterly about his guilt; (*d*) the subject is able to manage the situation, when a patient accidentally upsets the work area several times; (*e*) the subject does not feel that some patients who act in an unconventional manner are being "mean and nasty"; (*f*) the subject is not entertained by the entertaining conversation and overly friendly attention of an attractive, young male patient; and (*g*) the subject feels that the physical environment of the clinical service provides support to the student-patient communication. The evidence in this area of Attitudes and Activities of Patients indicate a pronounced change to favorable attitudes at the end of the affiliation.

In answer to the question, (*a*) were the attitudes changes increased favorable attitudes or decreased favorable attitudes, the data show that in all areas, except Personal Relations, there was an increase in favorable attitudes, at the end of the twelve-week affiliation. The answer to question (*b*) what statements rated and/or effected the greatest attitude changes in the areas of Personal Relations, Personal Feelings, and Attitudes and Activities of Patients, is demonstrated in the above discussion and will be indicated throughout the text. The greatest favorable changes appear to be a strengthening of healthy attitudes in the areas of Personal Feelings and Attitudes and Activities of Patients. The greatest unfavorable changes imply unhealthy attitudes at the outset or a weakening of desirable attitudes in the

area of Personal Relations. In the light of these data, the change suggested for future curriculum planning is earlier recognition of the strengths and weaknesses of attitudes of nursing students toward psychiatric nursing and toward the psychiatric affiliation.

In addition to the statistical analysis, an examination was made of the responses to each statement of the initial test and the re-test to find the per cent of favorable and unfavorable replies by the subjects. In interpretations, attention is drawn to the unfavorable responses because they indicate the negative attitudes student nurses have toward psychiatric nursing and toward the psychiatric affiliation. A third purpose of this study was to determine whether attitudes of students reveal implications relative to future program planning for the student during the psychiatric affiliation. Consequently, the following interpretations were made of the unfavorable responses to the attitude questionnaire.

### 1. *Preconceptions*

Responses to *Preconceptions* indicate that student nurses arrive at the psychiatric affiliation with a number of inaccurate preconceived beliefs, i.e., the use of padded cells to restrict very disturbed mentally ill patients, the fatal prognosis of mental illness, that abnormal behavior is foreign to normal behavior, that association with the mentally ill may cause mental illness, that the mental institution does not aid the individual patient, that psychotics should be hospitalized to restrict behavior for the sake of society but not for treatment of an illness. At the end of the affiliation these unfavorable responses were retained by only a small per cent of the students.

The highest per cent (71+) of unfavorable preconceptions centered about the statements: (a) the psychotic patient should be hospitalized for the sake of the safety of society; (b) the mental institution should have padded cells for the very bad patients; (c) current literature, T.V., and newspaper items high-lighting episodes of insanity tend to make people fear the psychotic patient; and (d) people of the community would not house or rent rooms in their homes to a former psychotic patient.

The low distribution ratings around the mean of *Preconceptions*, in Table 3, suggests that (a) preconceptions dominate the unfavorable attitudes of nursing students toward psychiatric nursing; (b) preconceptions reflect the negative attitudes of society, and/or the people with whom the students live and work, before entering the psychiatric affiliation.

### 2. *Expectations*

The greatest changes to favorable responses on the post-test in *Expectations* were centered about the statements: (a) certain patients are not wrongfully



institutionalized; (b) student sometimes finds she "has nothing to do" on the service; and (c) the special terminology used in psychiatry is not confusing.

The significant difference between the two means (Table 3) shows changes to favorable attitudes. This component discloses whether or not the subjects have favorable, appropriate, expectations with respect to the clinical practice service, the nurse leaders, the appearance of the mentally ill and their treatment and care, the role of the psychiatric nurse and the role of the student in psychiatric nursing, the psychiatric theory and instruction. The mean scores reflect the favorable responses in this area.

The highest per cent of unfavorable responses were centered about the statements: (a) the work is not explained sufficiently well; (b) the nurse in charge really does not like student nurses; (c) the work done by patients really should be done by employees, and; and (d) close contact with psychotic behavior is apt to change normal into abnormal behavior.

These data suggest that: (a) attitudes do change favorably during the twelve weeks; and (b) certain, specific, unfavorable attitudes are retained during the affiliation.

### 3. *Personal Relations*

In *Personal Relations*, the greatest unfavorable changes on the post-test were centered about the statements that: (a) the subject does not find relief from anxiety by talking over problems with her parents; (b) parents are unnecessarily concerned over the subject; (c) the subject does not feel at ease while working with the nursing personnel on the services; and (d) if some changes were made in the activities and atmosphere of the nurses' home and hospital, the subject would have had a happier experience here. Although the per cent of unfavorable ratings to these items is below the average response of the total group, attention is called to them since they seem to indicate feelings of insecurity and lack of social understanding.

The data in this category as indicated by per cent ratings suggest that: (a) attitudes in the area of personal relations do change to unfavorable ones during the twelve-week affiliation; and (b) unfavorable attitudes which are increased at the end of the twelve-week period indicate feelings of insecurity and lack of social understanding.

### 4. *Personal Feelings*

In *Personal Feelings*, the highest per cent of unfavorable responses on the initial test which were increased on the post-test, were centered about the statements that the subject: (a) felt shy, timid in the presence of the psychotic patient; (b) felt troubled over disliking a psychotic patient; (c)

had difficulty in covering up feelings of dislike for a certain patient; (d) worried a long time over an embarrassing incident that occurred on the ward; and (e) found excuse easily for failure to complete an assignment of class work or nursing service.

Data presented in Table 3 together with the per cent ratings in this area suggest that: (a) attitudes of personal feelings do change to favorable ones during the twelve-week affiliation; and (b) unfavorable attitudes which increased at the end of the twelve-week period indicate psychological insecurity and feelings of inadequacy toward psychiatric nursing.

### 5. *Attitudes and Activities of Patients*

The highest per cent of unfavorable responses on the initial and post-tests, centered about the statements that the subject: (a) is unable to continue a conversation with a patient who is not responding; (b) wanted to scold a patient who repeatedly soiled himself; (c) was afraid of the overactive patient; (d) was afraid of physical injury from the assaultive patient; (e) developed fear in anticipation of an attack, not physical injury; (f) felt impatient with elderly patients who demanded that certain things be done their way; and (g) rejected the patient who took a great deal of reassurance yet appeared dissatisfied with it.

Data presented in Table 3 plus the per cent ratings in this area suggest that: (a) favorable attitudes in the area of Attitudes and Activities of Patients do increase during the twelve-week period; and (b) the unfavorable attitudes which were retained to a large extent at the end of the twelve-week affiliation, indicate fear of and frustration by the psychotic behavior of the mentally ill patients.

## H. SUMMARY

The purpose of the thesis was to determine: (a) the expressed attitudes of thirty-five student nurses toward psychiatric nursing and toward the psychiatric affiliation; and (b) the revealed implications of these attitudes in improving future program planning for students during the psychiatric nursing period.

A psychiatric nursing attitude questionnaire of 100 statements was constructed to gauge the opinions of students in the areas of: (a) preconceptions; (b) expectations; (c) personal relations; (d) personal feelings; and (e) the attitudes and activities of patients with whom the student works. The questionnaire was administered in the following way: (a) Part I, Preconceptions, was given in individual oral conferences with each subject, during



the first few days of their affiliation and again during the final week of service; and (b) Part II, Expectations, Part III, Personal Relations, Part IV, Personal Feelings, and Part V, Attitudes and Activities of Patients, were given to all of the subjects as a written response in a group setting during the sixth week of affiliation, and again, during the twelfth week at the termination of the affiliation.

A statistical analysis was made of the responses of the total group of thirty-five subjects to the initial test and the post-test. A Pearson correlation, or  $r$ , was computed between the two sets of scores, ungrouped data, in a double entry table. The scores of the total group on the initial test ranged from 54 to 81. The scores of the total group on the post-test ranged from 54 to 87. The coefficient, or  $r$ , was .658. The obtained standard error or  $r$  was .17.

The  $t$  test of a difference between means, when the means are uncorrelated, was applied to mean scores on the five parts of the attitude questionnaire initial and post-tests to determine whether the difference in the means of the two tests was significant or whether it was probably due to chance. The  $t$  test revealed that the mean difference was statistically significant at the one per cent level in the areas of Personal Relations, Personal Feelings, and, Attitudes and Activities of Patients while in the areas of Preconceptions and Expectations it was at the five per cent and 10 per cent level, respectively. The mean scores were computed to determine changes or shifts in attitudes as expressed by differences in the first and final tests.

Further analysis of results was made by studying the per cent of frequency of various responses on each of the five parts in the questionnaire on the initial and post-tests. Thus the favorable and unfavorable attitudes and changes in attitudes toward the specific objects of the questionnaire were determined.

These data revealed that at the end of the affiliation there was an increase in favorable attitudes in all areas, except Personal Relations, in which there was an increase in unfavorable attitudes. Preconceptions, however, dominated the unfavorable attitudes in all other areas.

### I. CONCLUSIONS

Although no definite generalizations can be made in relation to the attitudes held by student nurses coming for psychiatric affiliation from schools of nursing in general, the findings of this study indicate:

1. Students arrive at the affiliation with dominant preconceptions toward the mentally ill patients, mental illness, and the treatment of mentally ill patients.

2. Preconceptions are developed from the opinions of classmates, who have had the psychiatric nursing experience, the beliefs of parents, friends, neighbors, and the community in general.

3. Students have preconceptions toward mentally ill patients, mental illness, and the treatment of the mentally ill patients, at the end of the affiliation, which indicate lack of knowledge, or resistance to knowledge of the cause, symptoms, prognosis, and treatment of mental illness.

4. Preconceptions predominate the unfavorable attitudes of students toward psychiatric nursing and toward the psychiatric affiliation.

5. Some students have inappropriate expectations concerning the work of the clinical practice service, the nurse leaders, the appearance of the mentally ill patients, the role of the psychiatric nurse, and the role of the student in psychiatric nursing.

6. Inappropriate expectations of students seem to be related to lack of experience, lack of pre-psychiatric affiliation orientation, and feelings of insecurity.

7. Some students who have negative attitudes in personal relations with normal people find these unfavorable responses increased at the end of the affiliation, which seems to indicate deep feelings of insecurity and lack of social understanding.

8. The initial negative attitudes of some students in the area of personal feelings changed markedly to favorable attitudes at the end of the affiliation.

9. Some students have negative attitudes of fear and frustration by the psychotic behavior of certain patients, in certain situations.

10. While the attitudes of student nurses toward psychiatric nursing and toward the psychiatric affiliation may be determined by an attitude questionnaire, at the end of the experience more than 10 per cent of the 35 students gave unfavorable responses to 78 of the statements.

Some implications of this study of a selected group of nursing students in future program planning for the psychiatric affiliation are that:

1. All students be given a questionnaire prior to and at the end of the affiliation to determine their attitudes toward psychiatric nursing and toward the psychiatric affiliation.

2. The curriculum be so planned and administered, and the student nurses be so guided throughout the basic program, as to assist them in changing unwholesome initial attitudes into desirable, healthy attitudes toward the psychiatric nursing experience.

3. Pre-affiliation orientation to the psychiatric program by the home school be given to all students participating in the affiliation.



4. Pre-psychiatric preparation by educational courses such as psychology, sociology, neurology, nursing of children, all related to the basic theory of psychiatry be given to all students participating in the affiliation.

5. Negative attitudes of students nurses toward psychiatric nursing and toward the psychiatric affiliation be recognized early by the faculty and efforts made to correct and/or modify them constructively.

6. More joint planning be carried on between the home schools and the school of psychiatric nursing to help identify and meet negative attitudes before students have the psychiatric affiliation and to devise ways of helping them build and maintain healthy, desirable attitudes toward patients throughout the psychiatric nursing experience.

#### REFERENCES

1. BLACK, K. Basic principles of nursing in psychiatric hospitals. *Hospital Management*, 1952, **64**, 81-91.
2. ———. Human relations content in the basic curriculum. *Nursing Research*, 1956, **5**, 4-17.
3. CAVAGLIERI, N. A new mental health program. *Am. J. of Nursing*, 1953, **53**, 212-13.
4. COGGINS, M. Feelings expressed by a group of nursing students related to the anticipation of their psychiatric nursing experience. Unpublished Master's Thesis, School of Nursing, Catholic Univ. of America, Wash., D.C., 1956.
5. GALLOWAY, B. A study of attitudes toward psychiatric nursing in a selected group of collegiate and non-collegiate students as revealed by certain tools. Unpublished Master's Thesis, Boston Univ. School of Nursing, Boston, 1953.
6. GARRETT, H. *Statistics in Psychology and Education*. New York: Longman, Green, 1947.
7. GUILFORD, J. *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill, 1956.
8. JOHNSON, E. The construction of a psychiatric nursing attitude scale and a study of the conception and attitudes of 274 student nurses as influenced by a psychiatric affiliation. Unpublished Master's Thesis, School of Nursing, Catholic Univ. of America, Wash., D.C., 1950.
9. KALKMAN, H. *Introduction to Psychiatric Nursing*. New York: Blakiston Division, McGraw-Hill, 1950.
10. KOOS, L. *The Questionnaire in Education*. New York: Macmillan, 1928.
11. MATHENEY, R. How does the psychiatric nursing affiliation fail? *Am. J. of Nursing*, 1951, **51**, 684-87.
12. PAYNE, S. *The Art of Asking Questions*. Princeton: Princeton Univ. Press, 1951.
13. PEPLAU, H. Utilizing themes in nursing situations. *Am. J. of Nursing*, 1954, **54**, 326.
14. REMMERS, H., & GAGE, N. *Educational Measurement and Education*. New York: Harper, 1955.
15. ROBINSON, A. Orientation begins at home. *Am. J. of Nursing*, 1952, **52**, 1091-92.
16. SAMHAMMER, J. Anxiety and the psychiatric nursing affiliation. Unpublished Master's Thesis, School of Nursing, Univ. Penn., Philadelphia, Penn., 1948.
17. SCHWARTZ, M., & SHOCKLEY, E. *The Nurse and the Mental Patient*. New York: Russel Sage Foundation, 1956.

18. STEELE, K., & MANFREDA, M. Psychiatric Nursing. Philadelphia: F. A. Davis Company, 1955.
19. THURSTONE, L., & CHAVE, E. The Measurement of Attitude. Chicago: The Univ. Chicago Press, 1929.
20. TIEGS, E., CLARK, W., & THORPE, L. California Tests of Personality Manual. Hollywood: California Test Bureau, 1953.
21. WALL, E. Student appraisal of a psychiatric affiliation. *Am. J. of Nursing*, 1948, 48, 401.
22. WEISS, O. Attitudes in Psychiatric Nursing Care. New York: Putnam, 1954.
23. WRIGHT, M. An investigation of attitudes about mental illness. Unpublished Master's Thesis, School of Nursing, Univ. Penn., Philadelphia, Penn., 1954.
24. ZILBOORG, G. Mind, Medicine, and Man. New York: Harcourt, Brace, 1943.

126 West Main Street  
Williamson, New York



## A NOTE ON THE VALIDITY OF THE BERNBERG HUMAN RELATIONS INVENTORY\*<sup>1</sup>

*Department of Psychology, Utah State University*

DAVID S. GORFEIN AND LARRY M. ANDERSON

### A. INTRODUCTION

Bernberg (3) reports a technique (The Human Relations Inventory, HRI), that is purported to measure social conformity. His instrument based on the "direction of perception technique" of attitude measurement (2) and the Allport J-curve hypothesis succeeded, in validation studies, in distinguishing between normals and known deviant populations (delinquents and criminals).

If the HRI does measure social conformity, might it not also distinguish between individual susceptibility to yield to a majority influence situations, similar to that reported by Asch (1) and in attitude change studies similar to that of Zimbardo (5).

### B. PROCEDURE

The present investigators, as a portion of an over-all research program concerned with conformity behavior, had the opportunity to administer the HRI to their Ss.

Study A was an attitude change study similar to that reported by Zimbardo (5). Eighty male undergraduates, registered in Introductory Psychology at Utah State University participated as volunteers. Ss served in groups of six (6) individuals. Their task was to rate a case history of a Juvenile Delinquent (5). Following the collection of their ratings one of the experimenters pretended to tally the Ss' opinion of the history and the "tally" was then returned to each S ("so as to keep everything together for purposes of IBM scoring" the Ss were told to give them a rational reason for the return of the tally). Each "tally" was falsified so that each S would see himself as a minority of one in holding a particular opinion as Most Acceptable. Following an interpolated activity Ss were asked under a pretense to re-rate the case history. Conformity was measured as the

\* Received in the Editorial Office on March 19, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> I wish to thank Dr. Raymond Bernberg and Psychometric Affiliates for permission to use the HRI in this research.

amount of change in the opinion an *S* marked Most Acceptable from the first to the second measurement.

Study B was a replication of the experiment reported by Asch (1) in which a stooge group gave falsified judgments of the equal line in what appeared to the *S* to be a task in visual discrimination. The critical *Ss* verbal judgment was preceded in every case by the report of three girls who served as stooges for the experimenters.<sup>2</sup>

Twenty-one female volunteers from Introductory Psychology classes at Utah State University served as *Ss* for this study.

### C. RESULTS

An independent item analysis of the Bernberg HRI led to the elimination of five items (Nos. 1, 3, 6, 11, 15) of the HRI as failing to meet Bernberg's criterion of "J" distribution for our population. Scores therefore are based on the 32 scored items. The lower the individual score the greater the conformity indicated, according to Bernberg.

TABLE 1  
RELATIONSHIP BETWEEN THE HUMAN RELATIONS INVENTORY  
AND ATTITUDE CHANGE (STUDY A)

	Direction of change			Mean change
	In direction of message	No change	Negative change "boomerang"	
HRI score				
At and below median	16	27	1	.60
Above median	12	19	5	.44

For Study A the HRI fails to distinguish between individuals who are influenced in the attitude change. There is a trend for *Ss*, who react to the communication by changing their opinion in the opposite direction, to score above the median of this group (in the non-conformist direction) on the HRI.

In examining the relationship between HRI scores and susceptibility to influence in the replication of the Asch study we do find significant predictive value for the HRI (see Table 2). There is a significant difference between low and high *Ss* on the Human Relations Inventory in whether they yield at all and in the mean number of trials in which yielding occurred.

A correlation coefficient between the HRI and the Marlowe-Crowne Social

<sup>2</sup> The authors are indebted to Alleda Clark and Gary White who served as experimenters in the Asch study and to Alma Carlisle who scored the HRI and patiently performed an item analysis of the instrument.



Desirability scale was insignificant ( $-.03$ ) for 52 Introductory Psychology students.

TABLE 2  
RELATIONSHIP BETWEEN THE HUMAN RELATIONS INVENTORY  
AND YIELDING TO GROUP INFLUENCE (STUDY B)

HRI score			Number who yield 1 or more times	Number who fail to yield	Mean trials yielded
	At and below median		7	4	3.55
Above median		2	8	.50	

$$\chi^2 = 4.07, 1df, p < .05; F_{1, 18} = 5.94, p < .05.$$

#### \* D. DISCUSSION

Results of our study leave the question of the validity of the Human Relations Inventory still unanswered. There seems to be some evidence for the predictive value of the instrument in the Asch-type study, but failure to get similar support in the attitude change study with the much larger sample leaves the issue in doubt.

It had been intuitively assumed by the experimenters that there would be a high social desirability component in the HRI since the instrument is supposed to measure conformity. This hypothesis was not confirmed for our population.

Finally we may note that for the population in our sample five of the original Bernberg items washed out in item analysis. Researchers using the HRI would do well to perform an item analysis before undertaking use of the instrument as a conformity measure.

#### E. SUMMARY

As part of a research program in conformity behavior the Bernberg Human Relations Inventory was administered to students being tested in conformity situations. The HRI which attempts to measure conformity employing a paper-and-pencil test receives partial support in terms of behavioral data. A measure of Social Desirability fails to correlate significantly with the HRI. In general our results suggest that in further research the instrument may be of value.

#### REFERENCES

1. ASCH, S. E. Effects of group pressure upon the modification and distortion of judgments. In E. E. Maccoby, T. Newcomb, and E. L. Hartley (Eds.), *Readings in Social Psychology*. New York: Holt, 1958.

2. BERNBERG, R. E. The direction of perception technique of attitude measurement. *Int. J. Opin. Res.*, 1951, **5**, 397-406.
3. ———. A measure of social conformity. *J. Psychol.*, 1955, **39**, 89-96.
4. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
5. ZIMBARDO, P. G. Involvement and communication discrepancy as determinants of opinion conformity. *J. Abnorm. & Soc. Psychol.*, 1960, **60**, 86-94.

*Department of Psychology*  
*Utah State University*  
*Logan, Utah*



## A STIMULUS-BRAIN FEEDBACK SYSTEM FOR EVALUATION OF ALERTNESS\*

*Veterans Administration Hospital, Bedford, Massachusetts*

THOMAS MULHOLLAND AND SYLVIA RUNNALS

### A. INTRODUCTION

Stimulus-brain response feedback loops are familiar (2, 4, 11, 13, 14). These loops have included a stimulus which was triggered at some pre-determined time or phase of the brain rhythm or they have generated stimulus frequencies which caused regenerative feedback, e.g., "photic driving" of the brain rhythm.

Feedback loops based on *suppression* of the brain rhythm by stimulation are much less familiar. One such loop used to generate a series of alerting responses to visual and auditory stimuli has been described (8). That system utilized the suppression or "blocking" of the *alpha* rhythm (8-13 c.p.s.) in response to a stimulus.<sup>1</sup>

The *alpha* rhythm is the prominent brain rhythm for a relaxed (not sleeping) human. With a sudden change of visual stimulation especially, synchronous *alpha* is suppressed or supplanted by faster, dysynchronized rhythms, i.e., an alerting response occurs. A feedback loop is created when the stimulus, which causes alerting, is automatically produced whenever *alpha* occurs; when *alpha* is then suppressed (or supplanted by faster rhythms) the stimulus is automatically removed; when *alpha* reoccurs the stimulus reoccurs, etc. Under these conditions a series of alerting responses is generated. This paper concerns the temporal variation of such series of alerting responses, and the application of the feedback technique to the study of "habituation," "attention," and individual characterizations of alertness.

### B. APPARATUS

The feedback loop was composed of (a) a Grass Model V Polygraph to pick up and amplify the brain rhythms; (b) a variable frequency, variable selectivity filter of the brain rhythms; (c) an amplifier of filter output; (d)

\* Received in the Editorial Office on March 21, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> That feedback loop as well as the improved version described here was originally suggested by W. Storm vanLeeuwen based on his previous study of a feedback loop (personal communication, 1959).

a relay controlled by (c) and controlling (e) a lamp; (f) a human subject. Figure 1 presents the flow diagram.

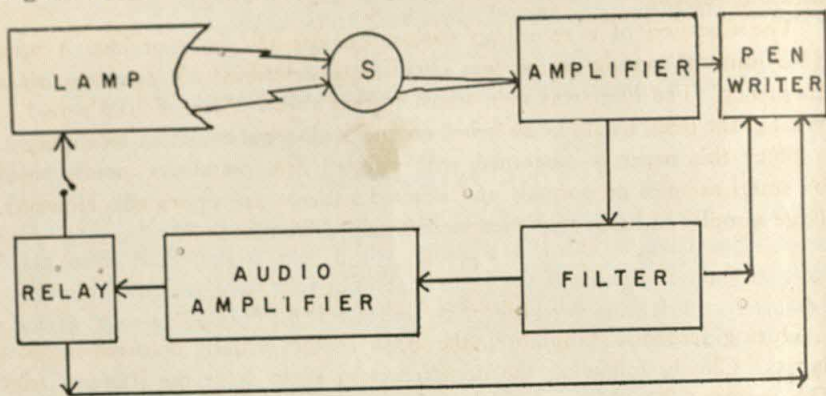


FIGURE 1  
FLOW DIAGRAM OF THE APPARATUS

The characteristics of the filter have been described previously (1). In order to obtain maximum output in the 8-12 c.p.s. band and to achieve adequate attenuation of frequencies greater than 13 c.p.s., a filter in the polygraph was also employed. This filter gave a 50 per cent or less response at frequencies greater than 15 c.p.s. The "Q" value (based upon maximum filter output equals 100 per cent) was approximately 5.8. This means that at a tuned frequency of 10 c.p.s. the frequencies yielding 70 per cent of maximum filter output or greater, ranged from 9.2 to 10.8 c.p.s. The time constant for the filter output to reach 63 per cent of maximum value at a tuned frequency of 10 c.p.s. was about .18 seconds. When the filter was tuned to 10 c.p.s., an input signal of 10 c.p.s. having an amplitude greater than 5-7  $\mu$ v would activate the stimulus relay. (The oscillographic record of the EEG reflected the attenuation by the high frequency filter in the polygraph which occurred prior to the write-out).

The visual stimulus was a 100 watt, incandescent lamp, located 10 feet from S. Incident illumination at the eyes was eight foot candles. The lamp could be switched *on* or *off* independently of the stimulus relay. This permitted an evaluation of system response in the absence of the feedback stimulus.

### C. PROCEDURE

Recordings were obtained in a sound reduced, air conditioned, dark room. The S was seated in a comfortable chair with his eyes closed. (The stimulus



was almost always presented through closed lids to prevent unnecessary tearing or irritation.) Bipolar scalp electrodes were placed over the parietal and occipital regions.

The duration of a recording session was one-half hour or less. A brief (1-2 minute) resting period was obtained to determine *S*'s dominant *alpha* frequency. The filter was then tuned to that value. After a brief period of resting, the lamp would be switched on and feedback stimulation would begin.

Since this paper is concerned with selected demonstrations, results based on small samples of normals and selected patients are presented. However, large samples of brain responses were obtained from each *S*.

## D. RESULTS

### 1. General System Behavior

During feedback stimulation, the *alpha* rhythm usually occurred in short bursts. Closely following the occurrences of *alpha* were the stimulus relay ON events. The intervals of *no-alpha* were followed by the relay OFF events. See Figure 2. The temporal pattern of successive ON and OFF events provided a detailed index of successive *alpha* and *no-alpha* events.

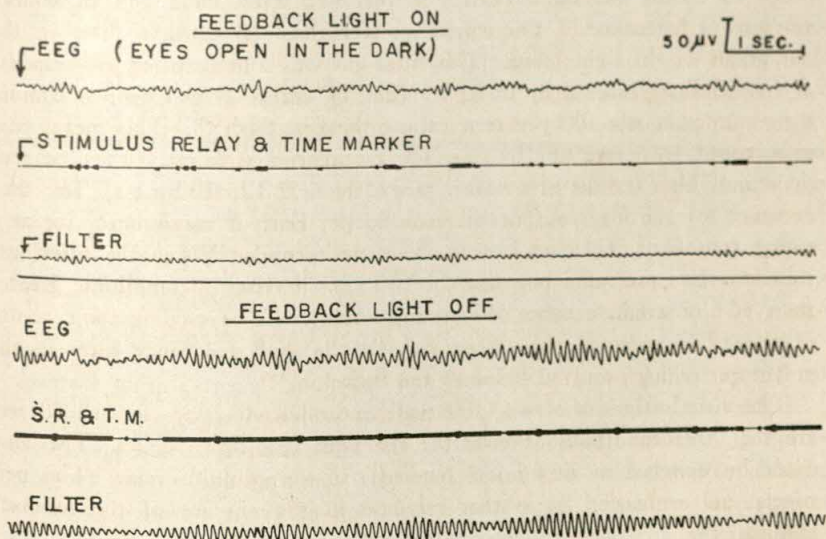


FIGURE 2

Example of system response and write-out. Record from a normal woman. The brief signal appearing at 1 sec. intervals is the time marker. Superimposed on the time marker are stimulus relay closures which are twice the amplitude of the time marker. Time lines on the EEG recording are curvilinear, producing an *apparent* .1 second lag of the relay marker relative to the EEG tracing.

## 2. Components of the Alerting Response

*a. Latency and duration.* The ON and OFF behavior of the lamp reflected both the latency and duration of successive alerting responses. In general, a reduced latency of *alpha* suppression would produce shorter *alpha* bursts and shorter stimulus-relay ON times. As the duration of the alerting response (alpha suppression or fast dysynchronous activity) increased, the longer the stimulus relay OFF times.

Since light ON times reflected *alpha* occurrence and OFF times reflected *no-alpha*, the terms "alpha" and "no-alpha" will be used in that sense for the remainder of this paper. The unit of measurement was .1 second.

*b. Fractionization of the alerting response.* By measuring successive durations at each serial position, series of event durations can be obtained for "alpha" and "no-alpha" events separately. As will be shown, the time series for "no-alpha" events was not a simple inversion of the series of "alpha" event durations.

(1) *Durations of successive "alpha" and "no-alpha" events during internal states of "attention."*

(a) *Procedure.* Six normal men were each tested under four different conditions of "attention." The conditions were: simple viewing—S was told to be aware of the light; counting by ones—S was told to count to himself the first 15 light stimuli, by ones; counting by three—S was told to count the first 15 light stimuli by threes, calling the first flash three, the next six, the next nine, etc.; counting by sixes—S was instructed to count the first 15 light stimuli by sixes, the first flash was six, the next 12, the next 18, etc. In this way the counting tasks varied in difficulty. Each S received the simple viewing conditions first and last in the experimental trials. The counting conditions were presented between the two simple viewing conditions. Each S received a different sequence of counting conditions.

The data analyzed were the durations of 10 "alpha" and "no-alpha" events just before feedback and 30 "alpha" and "no-alpha" events during feedback.

In the first analysis the two conditions of simple viewing were pooled as were the three conditions of counting the light stimuli. Using analysis of variance a comparison was made between: counting and simple viewing; subjects and sequence. Since the distribution of event durations was not Gaussian, the analysis of variance must be considered as suggestive not conclusive.

(b) *Results.* When feedback was initiated, "alpha" durations decreased and "no-alpha" durations increased relative to the immediately pre-



ceding resting condition (3, 5). As feedback continued, "alpha" durations increased slowly while "no-alpha" durations decreased rapidly, then more slowly (10). Comparison of event durations before feedback with those obtained during feedback showed the decrease of "alpha" durations was significant ( $F = 15.7, P < .001$ ) as was the increase of "no-alpha" durations ( $F = 21.6, P < .001$ ).

When  $S$  was counting the first 15 flashes a greater decrease of alpha durations occurred with feedback ( $F = 7.2, P < .01$ ) while a greater increase of "no-alpha" durations occurred ( $F = 24.0, P < .001$ ).

In a separate analysis of variance the three counting conditions were compared with each other. There was no evident change of the "alpha" durations as a function of the difficulty of the counting task. However, the average durations of "no-alpha" durations were shortest while counting by ones, more while counting by threes and longest while counting by sixes ( $F = 8.4, P < .001$ ). Non-parametric tests of this effect for each  $S$  separately gave significant results for only three of the six subjects.

In Figure 3, the mean "alpha" and "no-alpha" event durations during pooled counting conditions are presented. An average series of forty consecutive events is presented: ten just before the onset of feedback and thirty during the feedback stimulation. Each counting mean is based on 18 event durations; each no counting (simple viewing) mean was based on 12 event durations.

With the onset of feedback, "no-alpha" durations increased while "alpha" durations decreased. This effect was enhanced during the counting of flashes. In addition, a marked attenuation of the variance of average "alpha" durations during counting was seen. However, after the 15th flash the variance of the average "alpha" durations increased.

In a second experiment on this problem eight normal men were tested as in Experiment 1 under four different attention conditions: *A*, simple viewing; *B*, viewing and silently counting the first 15 flashes; *C*, viewing, counting and judging if the 15th flash were brighter or dimmer than the preceding one; *D*, viewing, counting and judging if the 20th flash were brighter or dimmer than the preceding flash. These conditions were presented in four different sequences using a Latin Square design.

In general the results were the same as in the previous experiment. With the onset of feedback "alpha" durations significantly decreased, "no-alpha" durations significantly increased.

This effect was not the same magnitude for each condition. In general longest "alpha" durations were obtained for the condition of simple viewing

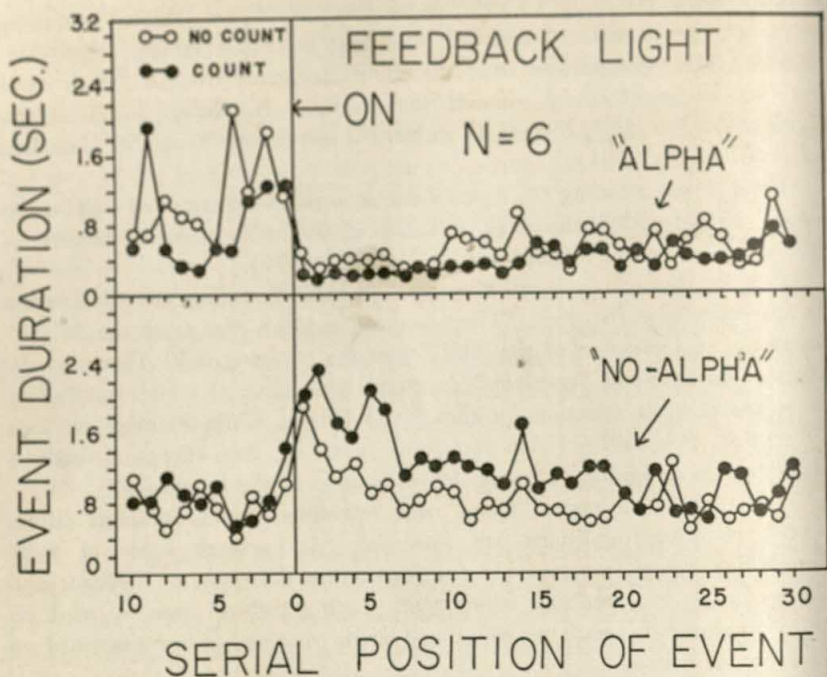


FIGURE 3  
MEAN EVENT DURATIONS "ALPHA" AND "NO-ALPHA" BEFORE AND DURING  
FEEDBACK STIMULATION UNDER TWO CONDITIONS OF ATTENTION

(A), and shorter for the other conditions. The reverse was true for "no-alpha" conditions. In addition the temporal changes of "no-alpha" durations during feedback varied significantly ( $F = 9.3$ ,  $P < .001$ ) as a function of the kind of counting condition present. Table 1 presents a summary of these results. Each score is based on a pooling of 5 successive events for the 8 subjects.

(2). *Other differential effects on "alpha" and "no-alpha" events.* In one study a patient receiving tranquilizing medication was studied over a period of 40 days. Of particular interest was the differential effect of feedback on the variance of successive "alpha" and "no-alpha" durations. In Figure 4, the transition from before to during feedback is shown. The mean durations of "alpha" and "no-alpha" events are based on the first seven days of recording. The general effects of feedback stimulation (increased "no-alpha" durations followed by decrease; decreased alpha-durations) are evident. The longer durations of "no-alpha" events before feedback reflect periods of



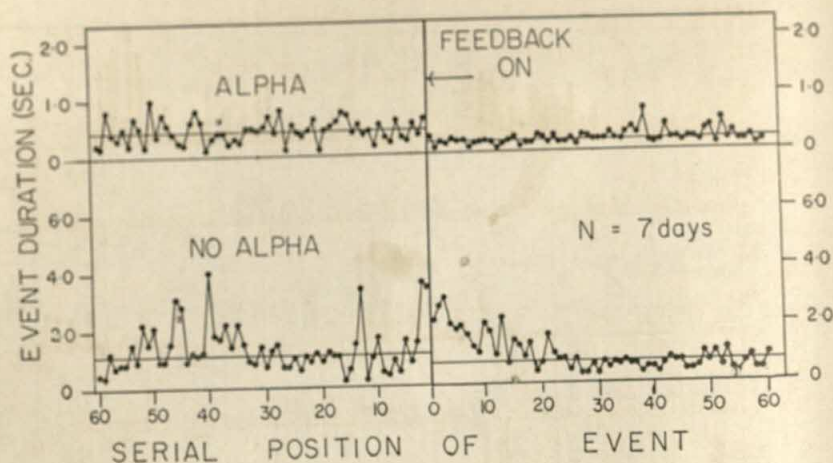


FIGURE 4  
MEAN EVENT DURATIONS FOR "ALPHA" AND "NO-ALPHA" BEFORE AND DURING  
FEEDBACK STIMULATION

slow (less than 8 c.p.s.) activity. The initial decrease of "alpha" durations is maintained during feedback while the "no-alpha" durations initially longer, decreased to an asymptote.

The marked reduction of the variance of "alpha" durations with feedback is shown. Some reduction of variance is also seen for the "no-alpha" durations. As the testing sessions continued this change of variation became less marked.

Another kind of differential change occurred when feedback stimulation influenced one component of the alerting response and not the other. In Figure 5 the detailed results for a normal man who was studied in the first experiment reported before are presented. This man showed a definite decrease in the pattern of "alpha" durations when feedback stimulation began but little or no change in the pattern of "no-alpha" durations. In contrast to this, another normal man from the same study showed both a decrease of "alpha" durations and an increase of "no-alpha" durations with the onset of feedback. See Figure 6.

(3). *Non-random variation of successive "alpha" and "no-alpha" response durations.* Some Ss during the first experiment described previously exhibited a non-random variation of successive response durations (6, 7, 8). Non-randomness was tested using a non-parametric runs test (9). Twenty-nine serial durations during feedback were designated "greater than" or

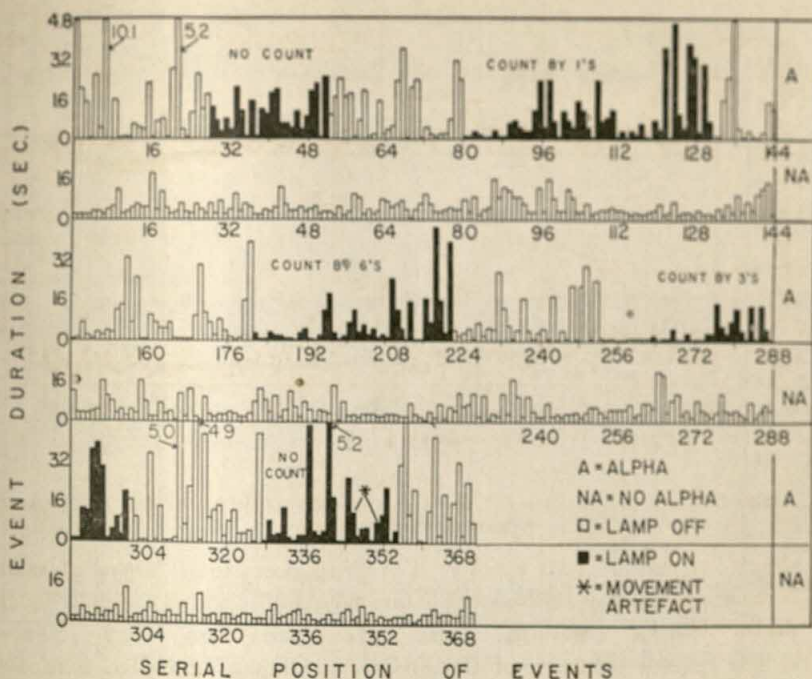


FIGURE 5

Successive "ALPHA" and "NO-ALPHA" event durations during conditions of simply viewing and counting flashes of feedback stimulation. Subject was a normal man showing a change only in "ALPHA" durations during feedback stimulation.

"less than" the previous duration. A runs test was made of the temporal grouping of these "greater than" and "less than" alternatives. In cases when a duration was equal to a previous duration it received the designation of the previous duration. In Table 2, the cases where non-random series were

TABLE 2  
NON-RANDOMNESS OF SERIAL ALPHA AND NO-ALPHA EVENT DURATIONS

Subjects	No-alpha					Alpha				
	No count		Counting by			No count		Counting by		
	I	II	1s	3s	6s	I	II	1s	3s	6s
1										
2			X							
3		X							X	X
4				X	X					
5		X						X		
6	X		X	X		X				

X = Significant at .05 level or less.



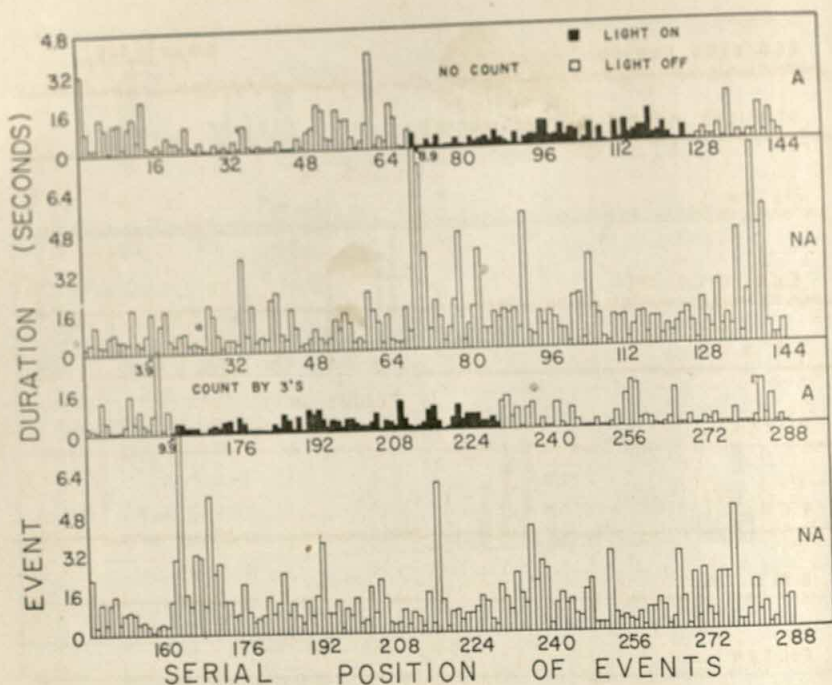


FIGURE 6

Successive "ALPHA" and "NO-ALPHA" event durations during conditions of simply viewing and counting flashes of feedback stimulation. Subject was a normal man showing a definite change in both "ALPHA" and "NO-ALPHA" durations during feedback stimulation.

inferred are shown. In each case the rate of alternation of the series as a function of serial position was greater than that expected by chance.

(4). *Hyper- and hypo-alerting patterns.* It would be useful to classify certain patterns of response as reflecting very little or very marked alerting. A minimal alerting would be characterized by long "alpha" bursts and short "no-alpha" periods (8). Figure 7 presents a recording obtained from a normal woman. With eyes closed she showed almost no response to feedback stimulation. With her eyes open, though a greater alerting was produced, the series of responses was indicative of a minimal alerting i.e., short "no-alpha" periods.

An example of a marked alerting to feedback is presented in Figure 8. The short "alpha" bursts and long "no-alpha" periods are clearly evident. This record was obtained from a patient having organic brain pathology. In

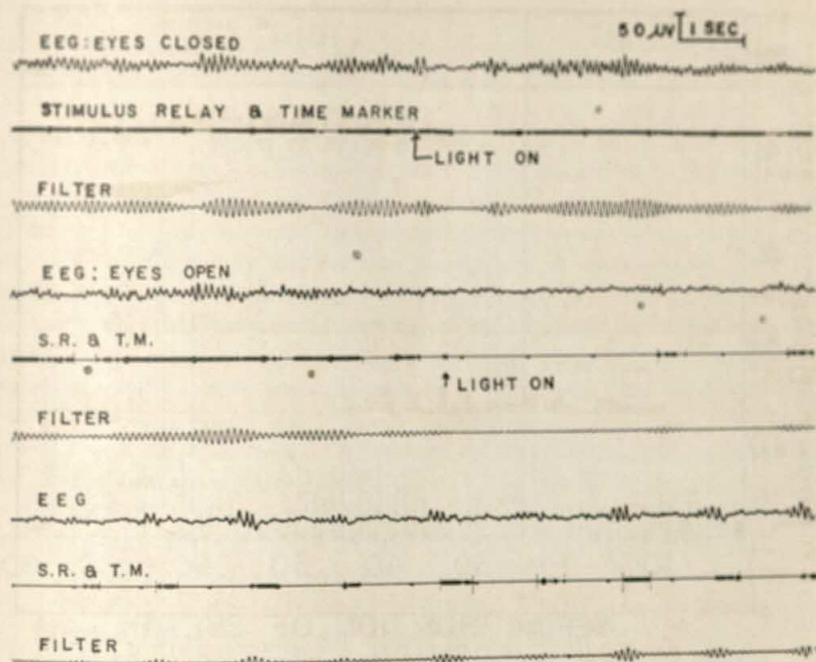


FIGURE 7  
MINIMAL ALERTING TO FEEDBACK STIMULATION. A NORMAL WOMAN

addition, the marked reduction of the variance of "alpha" durations relative to the variance of "no-alpha" durations is seen during feedback stimulation.

## E. DISCUSSION

### 1. Apparatus

The particular values chosen for filter selectivity ( $Q$ ) and stimulus relay threshold should depend on the particular needs of the experiment rather than upon some arbitrarily established value. In general a response to the dominant alpha  $\pm 2$  c.p.s., greater than  $2.3 \mu\text{v}$  should occur. The present inexpensive system approximates this ideal which can be realized by more critical filters and more sensitive relays.

Another problem concerns the individual variation of alpha amplitude. One may so adjust the system for each  $S$  that the stimulus relay is closed when a given fraction of  $S$ 's resting amplitude occurs or one can keep apparatus parameters fixed for all  $S$ s at the same values. The particular use



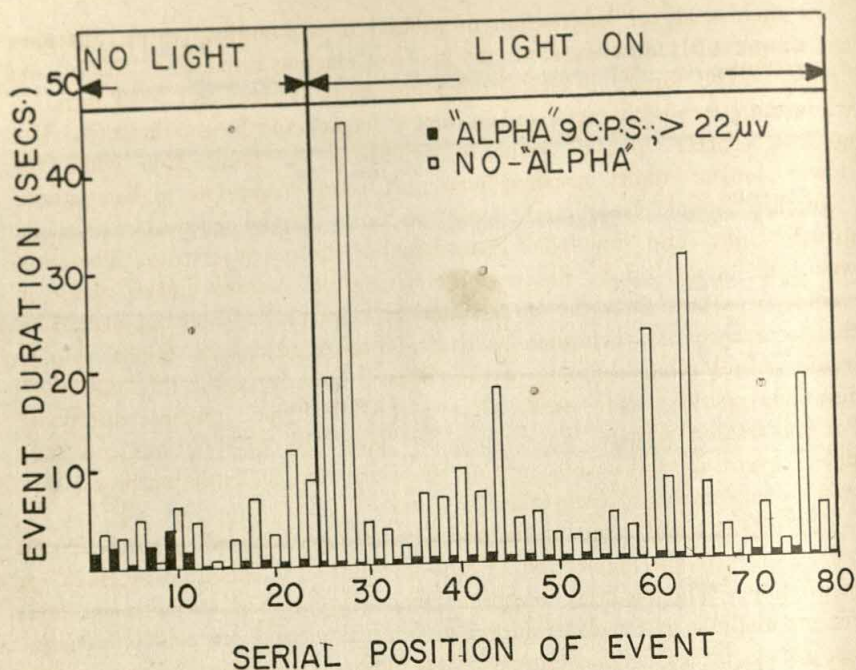


FIGURE 8  
HYPER-ALERTING TO FEEDBACK STIMULATION

of relative or absolute stimulus relay thresholds again must be related to the particular needs of the experiment.

## 2. Results

With feedback stimulation a series of "alpha" bursts and "no-alpha" intervals is generated. With the apparatus remaining stable, variations in the "alpha" burst durations reflect variations in the latency of the alerting response. The longer the latency, the longer the "alpha" burst duration. The variations of the "no-alpha" response durations reflect the variations in the persistence of the alerting response. The greater the duration the greater the persistence (by definition) of the alerting response.

During feedback stimulation the duration of "alpha" bursts (latency) was first brief and slowly increased. However, the duration of the "no-alpha" alerting response at first longer, decreased rapidly, then more slowly. These findings confirm those of Sharpless and Jasper (10, p. 660). In their terms these changes define habituation to the stimulus.

In their study of habituation of arousal to a stimulus (cats), Sharpless and Jasper (10) differentiated two kinds of alerting responses: "phasic" and "tonic." The former is characterized by short latency and brief duration while the latter is characterized by longer latency and longer duration. The quicker, shorter "phasic" response is resistant to habituation while the slower, longer "tonic" response was much more susceptible to habituation.

In this system "phasic" and "tonic" responses may be reflected respectively in the "alpha" and "no-alpha" components of the response series. The lamp would be turned off by the response having the *shortest* latency while it would go on only after the *longest* alerting response had been supplanted by the occurrence of synchronous activity. The durations of "alpha" bursts would sample relatively more of the "phasic" latencies while the "no-alpha" durations would sample some "phasic" and all "tonic" response durations. For this reason perhaps, the changes in "alpha" event durations and "no-alpha" event durations seen in this study are not evidently reciprocal nor simple complements of each other.

The efficient and rapid production of alerting responses by the feedback system suggests the application of this system to the quantitative study of habituation. The feedback technique may also provide a convenient and efficient analysis of the contribution of internal mental sets and operations, e.g., silent, internalized counting of the flashes yielded an increase in the durations and a decrease of the latencies of successive alerting responses relative to the conditions where no counting occurred. The feedback technique may find application in any studies where alerting or "attention" is an important variable, e.g., internal attention gradients; pharmacological agents; different emotional stimuli which may produce different amounts of alerting; reinforcing feedback stimuli, etc.

#### F. SUMMARY AND CONCLUSIONS

A simple electronic apparatus so filtered the EEG that selected frequencies in the *alpha* range (recorded from parietal-occipital locations) automatically caused a stimulus to occur. When alpha was then suppressed, the stimulus was automatically removed. When alpha reoccurred the stimulus automatically occurred again, etc. During this feedback stimulation the following phenomena were observed: (a) alpha tended to occur in a series of short "bursts" separated by periods of no-alpha activity; (b) the changes of duration of the alpha component and the no-alpha component over time were not necessarily inverse. Durations of alpha increased slowly while durations of no-alpha decreased rapidly and then decreased more slowly. In



some individuals the alpha bursts became shorter during feedback stimulation while no change was evident for the no-alpha durations. The variance of the alpha component showed a marked decrease while the variance of the no-alpha durations showed less decrease during feedback stimulation for some individuals; (c) for a small group of Ss tested here the temporal pattern of alpha and no-alpha response durations was significantly different for conditions of viewing and silently counting feedback flashes (more attention) compared to simply viewing the feedback stimulus. In general "alpha" durations were shorter, "no-alpha" durations longer and the variance of "alpha" durations was reduced during the period of internalized, silent counting of the stimulus; (d) the distributions of response durations were not Gaussian; (e) non-random variation of successive response durations sometimes occurred; (f) patterns of hypo- and hyper-alerting were described; (g) the hypothesis that the system fractionates "phasic" alerting response latencies from "tonic" and "phasic" alerting response durations was presented; and (h) some applications of the system to the study of "habituation" and "attention" and the effect of psychopharmacologic agents on these processes were suggested.

## REFERENCES

1. BECKER, H. D., MICKLE, W. A., & HEATH, R. G. A variable frequency variable selectivity filter for electroencephalography. *EEG Clin. Neurophysiol.*, 1958, **10**, 731-735.
2. CORRIOL, J., & GASTAUT, H. Le circuit déclancheur électronique de H. Shipton et W. Grey Walter. Son application en électroencéphalographie clinique. *Rev. Neurol.*, 1950, **82**, 608-613.
3. CRUIKSHANK, R. M. Human occipital brain potentials as affected by intensity-duration variables of visual stimulation. *J. Exp. Psychol.*, 1937, **21**, 625-641.
4. HEWLETT, M. G. T. An electronic trigger mechanism. *EEG Clin. Neurophysiol.*, 1951, **3**, 513-516.
5. JASPER, H. H., & CRUIKSHANK, R. M. Electroencephalography: II. Visual stimulation and the after image as affecting the occipital alpha rhythm. *J. Gen. Psychol.*, 1937, **17**, 29-48.
6. LARSSON, L. E. Sensitization of the startle blink and non-specific electroencephalographic response. *EEG Clin. Neurophysiol.*, 1960, **12**, 727-733.
7. MORRELL, L. K., & MORRELL, F. Periodic oscillation in the habituation curve of electrographic activation (abstract). *EEG Clin. Neurophysiol.*, 1960, **12**, 757.
8. MULHOLLAND, T., & RUNNALS, S. A brain response-sensory stimulus feedback system. Digest of Papers, 1961 International Conference on Medical Electronics, New York City. Publisher: Conference Committee for 1961 International Conference on Medical Electronics. L. E. Flory Chm., Princeton, N.J., 1961, p. 166.
9. SIEGEL, S. Non-parametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956, 52-58.
10. SHARPLESS, S., & JASPER, H. Habituation of the arousal reaction. *Brain*, 1956, **79**, 655-680.

11. SHIPTON, H. W. An electronic trigger circuit as an aid to physiological research. *J. Brit. I.R.E.*, 1949, **4**, 374-383.
12. TRAVIS, L. E., & HALL, M. E. Effect of visual after sensations upon brain potential patterning under different degrees of attention. *J. Exp. Psychol.*, 1938, **22**, 472-479.
13. TURTON, E. C. An electronic trigger used to assist in the EEG diagnosis of epilepsy. *EEG Clin. Neurophysiol.*, 1952, **1**, 83-91.
14. WALTER, V. J., & WALTER, W. G. The central effects of rhythmic sensory stimulation. *EEG Clin. Neurophysiol.*, 1949, **1**, 57-86.

*VA Hospital*

*200 Springs Road*

*Bedford, Massachusetts*



## BASIC MOTIVATION AND CONCEPT OF NURSING AS CHOSEN PROFESSION\*<sup>1</sup>

*Department of Psychology, Ohio State University, and Student Counseling Bureau,  
University of Minnesota*

EDWARD J. FURST, ALTON W. RAYGOR, AND ALAN P. CROFOOT

Of the many studies on personality and vocational choice, the great bulk have attempted to find some communality of traits in relation to career interest. These psychometric studies have been helpful, but unfortunately they tend to mask specific personality needs. In contrast, other studies have emphasized that specific personality needs, often unconscious, influence choice of curriculum or occupation (5, 7, 8). This hypothesis now seems well established. What remains to be explained is the fact that persons with dissimilar patterns of needs may choose the same vocation, and that persons with similar patterns of needs may choose different vocations. In a pioneer investigation, Dillon offered a theory to account for such discrepancies (2). She postulated that, in each case, the determining factor is the concept the person has of the profession he is considering and of himself in it, and the way this concept relates to his needs (2, p. 7). To test this theory, Dillon studied university students who had chosen teaching. She found basic motivation and concept to be closely related, and concluded that her findings supported well but did not prove the theory. Super (6, 7) has set forth a similar theory.

### A. PROBLEM, DATA REQUIRED, AND SAMPLE

#### 1. Problem

The present study was designed to test more thoroughly the hypothesis

\* Received in the Editorial Office on March 27, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was conducted while the authors were at the University of Michigan. Thanks are owed to the following: Frances H. Dillon for permission to adapt certain of her materials for this study; Haskell Cohen and George Witt for their interpretation of TAT stories; Suzanne L. Kallen for writing concept sketches; Richard L. Cutler, Zanwil Sperber, and Joan Williams for their matching and comparison of sketches; Educational Testing Service for permission to reproduce certain of their materials; the Nursing faculty of the University of Michigan for their cooperation; and finally, most especially the twenty young women who gave so generously of their time and effort.

Appendixes A-E, consisting of 9 pages of material, have been deposited with the American Documentation Institute. Order Document No. 7134, remitting \$1.25 for 35-mm. microfilm or \$1.25 for 6 by 8 in. photocopies.

offered by Dillon: the profession an individual chooses is the one that, according to his concept of it and as he imagines himself in it, seems to him to satisfy most fully the needs he feels the strongest pressure to fulfill (2, p. 6). This process is not necessarily conscious, for the individual may or may not be aware of the needs he feels pressure to satisfy. Nor is there any implication that the concept is necessarily realistic or clearly formulated.

Our initial plan required that we study a sample of students intending to enter various professional schools, so that we could obtain: (a) a clear picture of each individual's basic motivation; (b) each individual's concept of three professions, chosen to be rather contrasting in requirements and satisfactions; and (c) comparisons between each motivational picture and the respective concepts of the professions. These data would have enabled us to determine the relationship between order of preference for the three fields and degree to which personal needs seemed to have opportunity for satisfaction in the corresponding concepts. The plan also called for independent checks on the interpretation of the primary materials and for independent judgments of relationships, checks which Dillon necessarily had to omit in her study.

As promising as this plan seemed, two obstacles became readily apparent. One was the lack of suitable testing devices; and the other, the inability of many of our first group of subjects to formulate a clear concept of a second choice, let alone a third. Since the theory seemed promising, however, we decided to limit the study to one field—namely, nursing.

## 2. *Data Required*

The data required were of two kinds: first, a picture of basic motivation; and second, a picture of concept of the profession and of self in relation to it.

For a clear picture of motivation, it was felt necessary to include information on the individual's more specific desires and goals as well as on the fundamental needs more or less common in a given culture. It was also desirable to try to understand the ways in which the individual strives to fulfill her needs, her success or failure in achieving goals, and her means of compensation or other defense mechanisms; and furthermore, to pay particular attention to her relationships with other people. Since much of the data on motivation is contained in the individual's private world, which she may be reluctant or unable to admit, we decided to include projective techniques too.

Correspondingly, we felt that an analysis of each S's meanings concerning



nursing and herself in relation to it, to be valid, required data not only on conscious interpretations but also on the less conscious attitudes. The use of the expression, "concept of nursing," is thus somewhat imprecise.

### 3. Sample

The population consisted of student nurses. Twenty volunteers, representative of the freshman class in the University of Michigan School of Nursing, February, 1953 comprised the sample. As freshmen in a four-year degree program, they represented a naive group which had not yet had any formal study or work in nursing.

## B. PROCEDURE

### 1. Collection of Basic Data

Each *S* took part in five individual sessions spaced about a week apart. The second author was responsible for conducting these sessions, as outlined below.

*The TAT.* This technique included pictures 1, 2, 3GF, 4, 6GF, 8GF, 9GF, 12F, 13G, 13MF, 14, and 16 from the 1943 edition of the Murray Thematic Apperception Test. Each *S* was instructed to write out her stories, placing emphasis on plot.

An assistant then analyzed each set of stories and wrote an interpretive sketch. Although his analysis was aimed at certain common areas, it really amounted to a personality sketch rather than an assessment of needs. It was assumed that such a sketch would reveal more about motivation than an interpretation based solely upon needs and goals. As a check, another assistant read the stories and interpretations and made minor revisions.

*First Interview.* This first interview was aimed at bringing out dominant interests and values. The interviewer followed a schedule of questions that covered things in the immediate past and in the present which gave the *S* great satisfaction or dissatisfaction, what seemed important to her now, ways in which she would like to change her personality, etc. The interviewer did not follow these slavishly but used them as a guide. The interview was tape-recorded and transcribed.

*Interest Inventory.* As a supplement an interest inventory of the Like-Indifferent-Dislike variety was devised. The 156 items were drawn from two inventories (3) so as to sample relationships with family, relationships with the opposite sex, identification with others, interest in solitary activities, interest in leadership and social service activities, reactions to authority, autonomy-dependence, acceptance of impulses, severity with oneself, and preoccupation with cleanliness.

Each *S*'s choices were classified by these areas on a special summary sheet. This arrangement made it possible to note the relative emphasis upon L, I, and D in each area, and to study individual items and clusters of items, often across areas, to identify configurations. The following was one such configuration.

- (D) 6. Having a lot of close friends with whom I can talk about anything.
- (L) 36. Being very particular about the kind of people I have for friends.
- (I) 46. Chumming with someone from another group.
- (D) 56. Being friends with someone who thinks or feels very differently from the way I do.
- (L) 63. Doing things with a special group of intimate friends.
- (I) 67. Having long, intimate talks with one or more student friends.
- (I) 101. Studying with other students or working cooperatively with them on course projects.

Thus the foregoing choices suggested that this person preferred to narrow her circle of intimate friends to a very few and to avoid too intimate relations with others.

*TAT for Nurses.* To tap those unconscious attitudes and private perceptions of nursing, the authors developed a set of 10 pictures and drawings related to nursing (described in Appendix A). Each *S* wrote out her stories, again with emphasis upon plot.

*Questionnaire, Inventory, and Checklist.* These three instruments were designed to cover aspects of the concept of nursing. The questionnaire, which was relatively unstructured compared to the other two instruments, called for a page or half-page written response on various topics (Appendix B).

The inventory (Appendix C) consisted of 52 statements about nurses and nursing toward which the *S* was to record her opinion—Agree, Disagree, Uncertain. The check list (Appendix D) contained the same ideas but restated as possible reasons for wanting to be a nurse and rearranged in random order. Here she was to check all reasons that influenced her to choose nursing and to double check the five believed most influential. The responses to both instruments were later transferred to a summary sheet (Appendix E) consisting of a classification of the items. This permitted a rapid inspection of *S*'s opinions and reasons, as illustrated for case No. 8, given the fictitious name "Harriet."

Harriet's encircling of A after the first item below suggested that she expected nursing to call for the frequent use of initiative; her failure to check this as a major reason indicated that she did not regard herself as



*Challenge: independence*

Use initiative 25-9

Challenge to ability and ingenuity 5-34

Independent—little supervision 35-22

(A)	D	U
(A)	D	U ✓
A	(D)	U

motivated by this possibility. Her encircling of A after the next item suggested that she expected a challenge to her ability and ingenuity; her checking this on the reasons checklist indicated that this was one reason, though not one of the most important, for her choosing the field. Her encircling of D on the next item suggested that she did *not* expect nurses to have considerable freedom in planning and carrying out their own work; and her failure to check this as a reason indicated that she did not consider herself as going into nursing for the independent action it might offer.

*Second Interview.* This was intended mainly as a supplement to the projective and written instruments on nursing. Information was sought on the way each S pictured herself as a nurse, her notion of what an ideal nurse did in a day's work, ways in which she could measure up to that ideal, and ways in which she would fall short. This interview was also tape-recorded and transcribed.

## 2. Preparation of Motivational Sketches

The first main task in treating the basic data was to write a motivational sketch that would portray dominant needs and the ways in which the S was striving to satisfy them. In writing each sketch, the responsible assistant studied the TAT interpretation, the interview material, and the interest inventory responses. He did not refer to any of the material on nursing, with the exception of the second interview, in order to avoid confounding the motivational and nursing materials. It was assumed proper to glean additional points on motivation from the second interview, provided that these points did not enter into the concept of nursing. The assistant then organized these notes into a coherent descriptive sketch, one of which follows.

### MOTIVATIONAL SKETCH FOR CASE NO. 8, HARRIET

#### *Family relationships*

1. Harriet's needs for affection have had little opportunity for satisfaction in the home (TAT). Through over-control the parents have prevented the development of an ability to relate to others in an emotionally satisfying way (TAT, Int).

2. Harriet directs considerable hostility, some of it at the conscious level, toward parents (TAT, Int).

3. There has been little emancipation from this parental dominance. However, she has learned a variety of affection-gaining techniques in which she seeks the dominant role, and these have further hindered any satisfaction of her affectional needs (TAT, Int).

#### *Relationships with others*

1. Although strong needs for affection are indicated, other people are not usually seen as potential sources of love (TAT). Rather, they are regarded as rivals for whatever affection is obtainable in her few more intimate relationships (TAT, Int).

2. The need to dominate is present in this area as well (TAT, Int). As a result, the type of interpersonal relationship favored is an extremely close one in which this dominant role may be defined (Int).

3. The pattern of inventory responses strongly suggests a choosiness about making close friends, and a desire to narrow rather than widen the circle of intimate friends. This is consistent with her rejection of women's social groups and her relative indifference toward men (Int).

4. Harriet's hostility toward others comes out in the form of sarcasm and "cattiness"—habits which she openly admits and would like to change (Int).

#### *Heterosexual relationships*

1. Her expressed attitude toward the opposite sex is one of indifference: she "can get along without men." Although Harriet dates and participates in normal social activities, she shies away from emotional involvements (e.g., "going steady") (Int, Inv).

2. Yet at the unconscious level her needs for love and affection emerge strongly in this area (TAT).

3. The need to dominate, since it represents a technique in gaining this affection, appears very strongly here (TAT). This tendency has led to failure, and the resultant insecurity has led to considerable rationalization and autistic thinking concerning her relations with the opposite sex (Int).

4. Harriet desires marriage after a few years of work. Her picture of married life emphasizes being friendly rather than strict with children, and implies that she wishes to give the kind of affection that she herself did not receive (Int, Inv).

#### *Autonomy-Dependence*

1. Harriet shows a strong need for independence in decision-making, although she will turn to others for advice on "the little things" (Int, Inv).

2. She does not like to work under a dominant leader in small groups (Inv), nor to be told what to do when in a position of leadership (Int).

3. Her dislike of being "chairman of anything" strongly suggests that she wishes to avoid roles that make her dependent upon others (Int).



*Striving-Passivity*

1. The basic emotional attitude is active, directed toward striving rather than passivity (TAT, Int). Goals are seen as attainable through effort, and determinism of any kind is denied (Int).

2. However, some problems are attacked at the fantasy level. These typically involve the release of socially unacceptable behavior and are sado-masochistic in nature (TAT).

*Some desire to serve others, but with reservations*

1. In her work with spastic children, summers and currently with polio patients as a volunteer nurse's aid, Harriet emphasizes the satisfaction of seeing others improve from her efforts (Int) but she does not emphasize a desire to give nurture to others.

2. The TAT suggests that her need to serve others may be motivated, in part, by a sense of duty.

3. She dislikes campus or community social service activities (Int, Inv).

*Moderate need for recognition and self-esteem*

1. Grades serve not so much as an end in themselves but as a means of pleasing parents, emulating her brother, and satisfying herself that she can do it (Int).

*Other points*

1. Harriet shows a strong desire for physical activity (Int).

2. There is no evidence of strong intellectual needs or interests.

The motivational sketches were as complete as the sources of information permitted. If a sketch made no mention of a particular need, however, this was not taken to mean that the need was not present. Rather, it was assumed that the need was simply not evident (4).

*3. Preparation of Concept-of-Nursing Sketches*

The second main task in treating the basic data was to outline each S's concept of nursing and of herself in relation to it. In writing each sketch, the responsible assistant studied the interpretation of the TAT for Nurses, the free-response questionnaire, the material from the second interview, and a composite of the responses to the statements inventory and the reasons checklist. The sketch for case No. 8 illustrates form and scope:

## CONCEPT-OF-NURSING SKETCH FOR CASE NO. 8, HARRIET

*Role and characteristics of nurse*

1. Harriet emphasizes the nurse's role in curing physical ailments, and makes almost no mention of mental conditions.

2. Her concept of the nurse brings out the traditional virtues: patience, self-control, getting along with others, etc.

*Concept of patients and relationships with them*

1. As she sees it, nursing will satisfy her desires to serve others and to work with people.
2. Yet she wants to avoid much personal contact with patients. This stems partly from a fear of not receiving gratitude and affection from patients, and partly from a desire to escape boredom in bedside nursing. Surgical nursing, she believes, will permit her to avoid personal contact and boredom.
3. Although she sees patients as individuals, she generally sees them in a negative light—complaining, causing extra work, sulky, etc.

*Concept of other personnel and relationships with them*

1. Harriet wants pleasant associations and would like to meet interesting and important people in nursing. Fellow nurses she sees as possible friends and people to whom she can take problems.
2. However, she is very much concerned with the hierarchical relationships in the hospital and resents the doctors and head nurses as authority figures, although she feels they are available for help.
3. In her view, there are too many unfair rules for nurses and too much supervision. She believes that she will continually be told what to do and that this will hurt her sense of independence.

*Security of nursing*

Harriet seems to want security very much, and believes that nursing offers this. It is something to fall back on if circumstances should so require.

*Possibility of self-improvement*

Nursing offers a challenge to one's ability, she believes, and this she wants. In addition, she wants the preparation for marriage and family life which it offers.

*Other points*

1. Harriet feels that nursing carries considerable prestige, but that it would not improve her social standing.
2. She does not want to live at home after graduation and evidently feels quite strongly about this.
3. One of her reasons for entering nursing was to "repay the people who cured me" of polio.
4. Nursing offers variety and adventure—conditions she especially desires.

*4. Comparison of Sketches*

For each *S*, a motivational and a concept sketch were now available, and it was possible to treat these so as to test the hypothesis. The procedure was to compare these two sketches and judge the degree to which the concept



sketch seemed to provide an opportunity for the fulfillment of the needs. Accordingly, the authors independently rated each pair of sketches on a five-point scale: 5—very well; 4—reasonably well; 3—moderately well; 2—rather poorly; 1—hardly at all. The following comparison for case No. 8, illustrates the type of analysis each judge made, although these were done informally:

For Harriet, nursing provides an opportunity to serve others but in a very impersonal way. Surgical nursing would allow her to serve without having close personal contact with patients. Her desire to serve others seems to arise more from a sense of duty than from a compelling need to nurture others.

There is a remarkable consistency between the human relationships depicted in the motivational sketch and those in the concept of nursing. Coloring these relationships is an inability to relate to others in an emotionally satisfying way. The hostility toward other persons, the tendency to see them negatively and the failure to expect affection and gratitude emerge strongly in the view of nursing.

In other respects, too, there is a close correspondence. Nursing would permit the *S* to satisfy certain other needs. It would challenge her ability sufficiently, and so contribute to her self-esteem; win for her a degree of prestige and social approval; allow her to be physically active; provide a preparation for married life—which she strongly desires; and build an economic security for her to fall back on.

Yet the concept strongly suggests that certain basic needs are not likely to be met, or to be met with some difficulty. The *S*'s strong need to dominate others and to be independent do not seem to be fulfilled in those relationships in which she expects to find herself. Nor do her latent needs for affection find satisfaction in her imagined role as a nurse.

It would seem that the concept of nursing for this young woman permits, at best, a moderate satisfaction of her basic needs and desires. There are a few important discrepancies that would seriously limit need-fulfillment.

In order to check the ratings, we had three other judges match sketches. We selected five cases at random, removed all identifying marks from both sketches, randomized these sketches independently, and presented them individually to one staff member and two doctoral students in clinical psychology.

Further methods of treating the data consisted of correlation with other measures and qualitative interpretation of individual cases.

## C. RESULTS AND DISCUSSION

1. *Ratings of Need-fulfillment*

The mean ratings given by the original judges were 4.05, 3.85, and 3.50, respectively, with an over-all mean of 3.80. The corresponding standard deviations were .97, 1.27, and 1.12.

When the ratings of the judges were intercorrelated, the respective *rs* turned out to be .73, .85, and .86, with a mean of .81. Each of these was statistically significant at the .001 level. These data indicate that the judges agreed rather closely in their ratings of the extent to which the concept fulfilled the needs.

Further evidence of objectivity came from the judgments made by the three clinical psychologists. In the blind matching, these judges made eleven correct matches out of a possible fifteen. Since the chance probability of this occurrence is less than .001 (1), we concluded that the kinds of data in the sketches do differentiate between individuals and do contain "psychologically real" cues to which the trained observer can respond. After the matching, each of these judges also independently rated the five cases on need-fulfillment. These ratings were averaged for each case and the ranks of the average ratings correlated with those of the original judges. This resulted in a rank-difference correlation of .87, which also indicates high objectivity.

The statistic most relevant to the hypothesis was the over-all mean rating. We have concluded that the mean of 3.80 supports the hypothesis. This value indicates that the concept of nursing fulfilled the needs about "reasonably well." In other words, individuals who have chosen nursing but have not yet entered it, do tend to form a concept of the profession, and of themselves in it, which seems to them to best fulfill their strongest needs.

Both the present study and that of Dillon (2) did provide an additional type of evidence. This was the association between the ratings of need-fulfillment and relative certainty of choice of occupation. If the hypothesis is sound, these two variables should show correlation. Thus, those whose concepts show adequate opportunity for need-satisfaction would be expected to be relatively more certain of their choice than would those whose concepts show limited opportunity. This would imply, too, that those who are not too sure of their choice are considering other occupations as being more likely to satisfy their needs. To obtain the necessary statistics, we prepared a contingency table in which average ratings of need-fulfillment were reduced to three categories—high, moderate, low—as were ratings of certainty of choice. The latter were based upon recorded utterances of the *Ss*. We then



computed chi square and contingency coefficient for these data as well as for the corresponding table in Dillon (2, p. 258), which involved 25 prospective teachers. Our data yielded a contingency coefficient of .56, with a theoretical maximum of .816, and a  $P$  of .06 for chi square. For the other study, the values were .76, .816, and .001, respectively. Taken together, these data offer strong support for the hypothesis.

One of the other relationships examined was that between the ratings of need-fulfillment and graduation and entry into nursing. A contingency table portraying these variables failed to reveal any significant relationships for this sample. It must be recognized that in the interim period of over three years many changes took place in these Ss. Eight of the 20 married and doubtless developed commitments to continue even if the choice of nursing was not the best, and for the sample as a whole, the basic motivations and concepts must surely have undergone some change. It must be recalled, too, that the concepts of nursing were not assumed to be realistic, and therefore would allow for persons with rather wishful concepts to persist in the educational program.

The relationship of the ratings to a measure of general ability is also of some interest. The coefficient of correlation between the first variable and the American Council on Education Psychological Examination (ACEPE) total scores turned out to be  $-.12$ . This value does not differ significantly from a zero  $r$ . Hence the highly intelligent were no more likely than those of more limited ability to show close agreement between sketches.

## 2. Qualitative Data

The statistical results cannot bring out the vivid relationships between personality and concept of self in the chosen profession. It remains for the case materials to speak for themselves, as they have for case No. 8. To demonstrate these often striking relationships, we have reproduced brief sketches on six other Ss. Three of these showed the highest degree of need-fulfillment, and three the lowest.

The first sketches are those of three Ss in whom the desire to serve others was quite strong and permeated their concepts of themselves as nurses. Beyond this similarity, important differences stood out. (The numbers after each pseudonym indicate the ratings given and the percentile rank on the ACEPE in the entire freshman class.)

*Case 11. Karen (5, 5, 5, 28).* The dominant characteristic of Karen's personality seems to be a passive dependency arising from over-control by her parents and from a strong feeling of obligation toward them for sacrifices

in her behalf. This dependency has led to a strong desire to please her parents, to considerable suppressed hostility toward them, and to a passive emotional attitude. This pattern extends to her relations with others. Karen continually attempts to do what she thinks will please others, in an effort to gain love and acceptance. Being of service and aid to others is her way of gaining belongingness and affection.

These characteristics show up clearly in Karen's concept of nursing. Her anticipated role of the nurse as comforter and counselor is extremely nurturant. Her needs for belongingness and dependency are met still further through warm relationships with superiors and co-workers, and through the security which nursing offers.

*Case 13, Mary (4, 5, 5. 95).* Mary's relations with her parents have been satisfying and free from tension although never particularly intense. The values of her parents have been so completely accepted that they now function as her own. In fact, the behavior that she builds on this foundation is active and independent.

Mary has considerable insight into her personality and tries to find social situations which best fit her needs. As she feels she is quick-tempered and emotional, she seeks out structured situations in which the possibility of such behavior is minimized.

An important part of Mary's value structure is her religious belief and the internalization of the Christian ethic. Other persons have value in their own right; love and respect form the basis for valid human relations. In addition, Mary feels a strong need to repay her parents for their sacrifices, a need now generalized into a desire to help others. Yet for her, there is one goal before all others: a warm and satisfying family environment of her own.

Mary's desire to help others is well met in her view of nursing. This service she sees as the nurse's prime function, but especially in the maintenance of the patient's mental health. She anticipates no conflict with authority figures in a hospital setting and expects no interference with her needs for independence. The spirit of the hospital is in fact that of cooperation between differentially trained equals. Finally, nursing is a fine preparation for marriage and family life as it provides valuable skills as well as general self-improvement.

*Case 16. Peggy (5, 5, 5. 40).* Perhaps more than any of the other Ss, Peggy presents a picture of the liberal humanitarian. Coming from a family characterized by warmth, affection, and respect, she finds her greatest satisfactions in sharing such relations with others. Her dominant



need is to associate with people and to share their problems. For her, marriage fulfills this need best.

Peggy's view of nursing reflects well her interest in other people and their welfare. Her urge to exercise authority as a nurse would seem to be consistent with the protective, nurturant attitude which she takes toward patients. Peggy sees the nurse as possessing high status and being respected by others, while at the same time respecting her patients and reacting to them fully as individuals. The hierarchical structure of the hospital carries no threat to her need for independence, for she accepts it and sees superiors as understanding and helpful. Nursing, for her, is an interim pursuit which can prepare her well for marriage.

In the three other sketches, there was consistency between basic motivation and concept but not strong evidence of need-fulfillment. Choice of nursing did not seem to reflect any genuine motives for the activity.

*Case 19. Ruth (2, 1, 1. 62).* Faced with a threatening world, Ruth lives in a continual state of fear. Unfortunately, she has few techniques to deal with her fears. Even intense religious convictions have been of little aid in changing her view of the world.

Ruth's home life has been extremely unsatisfying and she has an intense hostility to both parents whom she sees as having only punished and deserted her. So, in spite of strong needs for affection, she is unable to allow herself to become involved in relationships where affection becomes available, for fear that desertion and pain will result.

Although she consciously craves independence, Ruth has been unable to break away from dependence on others. All her goals are organized around a dependent and passive security.

Ruth's concept of nursing offers little hope of satisfying her needs. There are no satisfactions to be obtained from either the work or the associations that may be formed. Even service to patients is motivated by an abstract concept of "duty" related to a need for atonement because of her hostility to others. Moreover, although the nurse is fully dependent on doctors and other superiors, she is always threatened by them and can expect little but blame and criticism. Nor does nursing provide in itself the security she seeks, for in her view there will be continual competition and humiliation.

*Case 20. Sandra (2, 2, 2. 83).* Sandra is interesting in that her concept of nursing seems to be built around her idealized self rather than any realistic picture. Typically she is a dependent, passive girl who has been unhappy at home and, in an effort to resolve the guilt resulting from her hatred of her parents, has turned most of her aggressive impulses against

herself. As a result we find a very negative self-picture and a rich fantasy life filled with self-punitive themes. Beneath this, however, is a continual striving for affection and security which she feels may best be attained in marriage.

Although the nurse Sandra pictures is her own idealized self—charming, warm, and nurturant, the situation in which this ideal operates provides little satisfaction for either her ideal or real self. She is resentful of the authority structure, frightened by superiors, and distrustful and critical of her fellow nurses. Even her security goals are not realized in nursing, for the job itself provides no permanent rewards nor does nursing increase the probability of her finding the desired strong, stable marriage partner.

*Case 4. (3, 1, 2. 42).* Unusually strong needs for self-esteem, social approval, and social status dominate Dorothy's personality. Good grades, a college education, a brief career—all are important because they give her a feeling of amounting to something and they win for her the approval of others, especially parents. The need for affiliation seems strong, but the emphasis is on what can be gained from others rather than what can be done to enrich their lives. Correspondingly, while the desire for marriage is strong, what Dorothy primarily seeks here are the personal advantages of social status and financial security. (In a follow-up interview, she exclaimed with joy: "I got my 'med' student!")

That Dorothy should prove to be an aggressive, striving person is not at all surprising. That she should show ambivalence toward independence is surprising, however, until one notes the high degree of maternal control to which she has been subjected and from which she is slowly freeing herself.

Dorothy's concept of nursing seems rather inconsistent with her motivation. On one side, the motivational material does not support the emphasis placed upon nurturant service to others. On the other, the concept provides no outlet for the strong status needs. However, the ambivalence toward independence does show up well in the concept; there is a strong resentment toward authority and supervision, yet at the same time a desire to seek help and structure in the nursing situation.

### 3. Developmental Factors

Although it was not our main purpose to analyze development, the material contains some intriguing leads. In many cases we were able to detect some close relations between present behavior and earlier experiences, particularly those involving family interaction. In cases where there was



strong evidence of lack of warmth and emotional expression in the family, we tended to find evidence of a lack of empathy and an emotional constriction in relating to others. In some instances this carried over into a preference for a formal and socially distant professional role. The finding that eight out of 20 *Ss* showed a decided deprivation of affect led us to speculate that compensation might be operating. Could it be that many women unconsciously choose a professional role such as nursing in order to give others the love and security which was denied them in their upbringing? It is our conviction that this mechanism was operating in at least a few of the *Ss*, but by no means in all showing deprivation.

In several cases there was evidence of warm and satisfying family relationships, together with high empathy and warmth toward others in the imagined role of nurse. These were obviously not instances of compensation, but of the persistence of long-established patterns of behavior into other areas of life.

More than any other *S* Ruth did appear to be compensating for loss of love. Figure 1 brings out the main influences that appear to have shaped her personality and eventual choice of nursing.

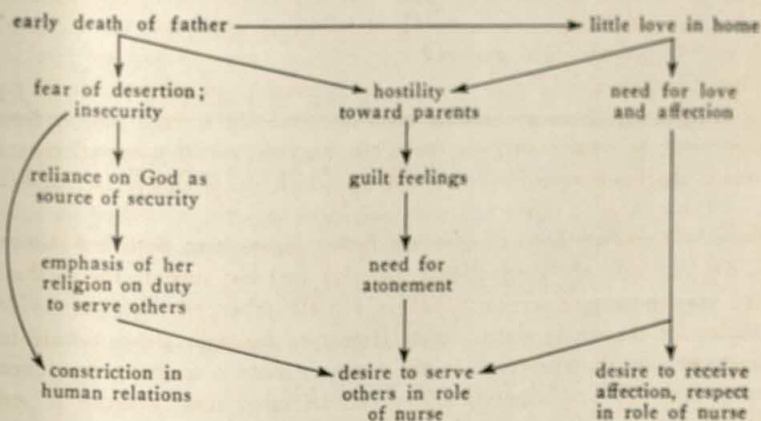


FIGURE 1  
IMPORTANT FACTORS IN RUTH'S DEVELOPMENT

As a final illustration, we refer again to case No. 8, Harriet. It is noteworthy that her recovery from polio in childhood generated a desire to enter nursing in order to repay society. This re-payment motive seemed to account, more than anything else, for this *S* being in nursing. Other develop-

mental and personality factors, however, have shaped her anticipated role. Their influence can readily be seen in the case materials presented earlier.

#### D. SUMMARY

This study purported to test the hypothesis that the profession an individual chooses is the one which, according to his concept of it and as he imagines himself in it, seems to him to satisfy most fully his strongest needs. The essential operation was to determine the relationship between (a) basic motivation; and (b) concept of chosen profession and of self in relation to this profession. A motivational and a concept sketch were accordingly outlined for 20 student nurses. The corresponding sketches were then compared and the concept sketch rated on need-fulfillment. The mean over-all rating of 3.80 on a 5-point scale indicated that the concept of nursing fulfilled the needs about "reasonably well." Further support for the hypothesis came from the correlation between the ratings and relative certainty of choice of profession. Descriptive material was presented for seven cases to illustrate specific relationships between personality and concept of self in the chosen profession, and to explore important developmental influences.

#### REFERENCES

1. CHAPMAN, D. W. The generalized problem of correct matchings. *Ann. Math. Statist.*, 1935, **6**, 85-95.
2. DILLON, F. H. The Relationship between Basic Motivation and Choice of Teaching as a Profession. Unpublished doctoral dissertation, Univ. Chicago, 1949.
3. Educational Testing Service. PEA Interests and Activities Index Forms 8.2b and 8.2c; Inventory of Personal-Social Relationships, Part I. Princeton: Author, 1939, 1950.
4. MASLOW, A. H. A theory of human motivation. *Psychol. Rev.*, 1943, **50**, 370-396.
5. MURPHY, L. B., & LADD, H. Emotional Factors in Learning. New York: Columbia Univ. Press, 1944. Pp. 404.
6. SUPER, D. E. Vocational adjustment: implementing a self-concept. *Occupations*, 1951, **30**, 88-92.
7. ———. The Psychology of Careers. New York: Harper, 1957. Pp. 362.
8. VERNON, M. D. The drives which determine the choice of a career. *Brit. J. Educ. Psychol.*, 1937, **7**, 302-316; 1938, **8**, 1-15.

Department of Psychology  
The Ohio State University  
Columbus 10, Ohio



## DETECTION OF DIFFERENCES IN DURATION OF ACOUSTIC AND ELECTRICAL CUTANEOUS STIMULI IN A VIGILANCE TASK\*<sup>1</sup>

*U.S. Army Medical Research Laboratory, Fort Knox, Kentucky and  
Lockheed-Georgia Company, Marietta, Georgia*

---

M. LOEB AND G. R. HAWKES

---

### A. INTRODUCTION

A considerable body of research has been performed on vigilance or monitoring behavior (1). Generally, when observers are to detect obscure signals—especially transient changes in physical stimuli—detection probability declines as a function of time on task. The majority of studies in this area have employed visual displays.

There are practical and theoretical reasons for performing research using non-visual displays. In the operational situation the visual modality is frequently pre-empted by task requirements. (This is true, to a lesser extent for the auditory modality.) Since perceptual capacities are not precisely parallel in different modalities, obtained differences in vigilance for comparable signals in different modalities should provide insight into the nature of the mechanism of attention.

Vigilance tasks may be arbitrarily divided into two general categories—those requiring detection of the presence of signals and those requiring detection of changes in signals.

The authors have recently published a series of studies (6, 7, 10) concerned with detection of the presence of brief, random noise signals in the acoustic, electrical cutaneous, or mechanical vibratory cutaneous channels. Operators were able to detect simple auditory signals with a high degree of efficiency and detection probability remained high throughout the session. These results are consistent with those reported by Elliott (5) and by Martz and Harris (12). On the other hand, detection probability for faint cutaneous signals was significantly lower, and decreased with time

---

\* Received in the Editorial Office on March 28, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was aided by a contract between the Office of The Surgeon General, U.S. Army, and the University of Louisville. The authors are indebted to Dr. Ray Bixler for administrative aid and advice and to Mr. Don Kaufman, Miss Caroline Roulston, Mr. E. A. Schmidt, and Mr. Charles Gettys for technical assistance.

on task. In addition, a progressive tendency toward an increase in response time was noted. For moderately intense cutaneous signals, however, no decremental tendency was observed, either in terms of latency or of failures of detection.

Frequently as in such operational situations as the monitoring of sonar displays, the inspection of manufactured items, etc., observers are asked to detect changes in stimuli, rather than the simple presence of such stimuli. Several studies of detection of changes in auditory stimuli have been reported. Solandt and Partridge (13) reported a progressive decline in ability to discriminate changes in auditory pitch. Mackworth (11) reported a progressive decline in ability to detect increases in durations of pure tones. In addition, Buckner, Harabedian, and McGrath (3) have demonstrated a progressive decline in detection of auditory intensity differences.

A very recent experiment by Buckner and McGrath (4) published after initiation of this investigation, employed a design similar to that of the present study. In separate sessions observers monitored intensity changes in visual stimuli, auditory stimuli, simultaneous visual-auditory stimuli, auditory and visual stimuli (not simultaneous), or auditory, visual and simultaneous auditory-visual stimuli. Detection of auditory signals was superior to detection of visual ones. Simultaneous auditory and visual signals were detected better than either auditory or visual signals alone. When observers were monitoring signals presented in different modalities independently in the same session, detection probability for easily detected signals increased, that for less easily detected signals decreased, and the over-all performance was roughly equal to the mean.

Possible differences in detection of temporal changes in auditory and cutaneous signals have not been reported. It has been demonstrated that at moderately high intensities the detection of the mere presence of auditory stimuli was approximately equal to that for equally intense cutaneous signals (7). It also has been shown that for alerted conditions, temporal discrimination for auditory signals does not differ from that for cutaneous signals (8). The object of the experiment was to determine the efficiency with which changes in duration of auditory and cutaneous signals could be detected and the efficiency with which the detection could be maintained over an appreciable period of time.

#### B. APPARATUS AND PROCEDURE

Unless otherwise specified, the apparatus and procedure were identical to those of previous studies (6, 7, 10). Auditory stimuli were produced by feeding the output of a random-noise generator through an attenuator into



Permoflux PDR-8 earphones. Electrical cutaneous stimulation was produced by feeding the output of the random-noise generator through an amplifier into a 12 mm. diameter active electrode taped to the index finger pad and a 25 mm. diameter inactive electrode taped to the palm of the hand.

Each of the 24 *O*s (18 male and six female university students) participated in four sessions. In one session (condition "C") *O* was given only electrical cutaneous stimulation; in another only auditory signals. In a third (condition "CA"), *O*s received both cutaneous and auditory stimuli presented simultaneously. In a fourth session (condition "X") *O* was given either cutaneous or auditory stimulation, the kind of stimulation to be given at a particular instant being determined by a random selector switch. Sessions were presented in a counterbalanced order. None of the *O*s had previously participated in this type of experimentation.

The intensity of the electrical cutaneous signal was 5.1 db *SL* (i.e., above absolute threshold); in sessions utilizing auditory signals a matching procedure was employed to establish a signal strength subjectively equal to the cutaneous stimuli. The intensity of the auditory signals averaged about 45 db. It was shown in an earlier experiment (7) that detection of the presence of auditory and cutaneous signals of this intensity is good and is well maintained. It has been demonstrated (14) that cross-modality matches of this kind may be made reliably and that for scaling purposes such matches are valid ones. The cutaneous absolute threshold (*RL*) determination and the cutaneous and auditory match were made at the beginning of each session and repeated at the end of the session to ascertain whether or not the repeated stimulation of the session had had a different effect upon the two stimulus modalities.

Signals of 0.5 seconds suprathreshold duration (with an onset and decay time of 50 msec.) were presented regularly throughout the session, with a 2.0 second pause between stimuli; these stimuli were referred to as "distractors." Occasionally a "cue" stimulus of 1.0 second duration was presented, and *O*'s task was to press a telegraph key as rapidly as possible when he detected the occurrence of the "cue" stimulus. These durations (0.1 and 1.0 seconds) are readily identified under alerted conditions (8). The intervals between presentation of "cue" stimuli were 75, 112, 150, 187, 225, 262, or 300 seconds. There were four successive blocks of seven trials each, with each of the intervals used once (in a random order) in each block. Duration of a session was approximately 100 min.

Earphones and electrodes were worn by all *O*s during each session. Earphones were relatively comfortable and were mounted in Willson Sound Barrier earmuffs, providing 15-25 db attenuation of ambient noise (depend-

ent upon frequency). Ambient noise level in the experimental chamber was about 25 db. Actions of *O* were monitored with the aid of a closed-circuit TV system, and voice communication between *O* and *E* was possible. The *O*s were paid a standard amount for their participation, and in addition a cash bonus was offered for the best performance during the four experimental sessions.

Reaction time to the nearest msec was recorded from a Standard Electric timer. Failure to respond to a "cue" signal within 3.0 seconds after its initiation was recorded as an error of omission. All responses made before or after such 3.0 second periods were scored as false responses.

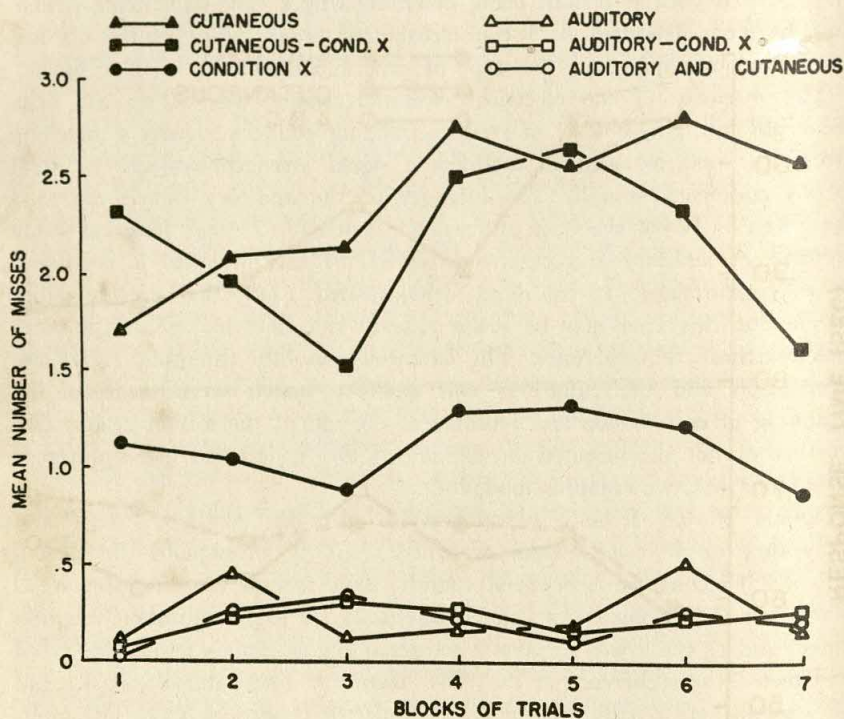


FIGURE 1  
ERRORS OF OMISSION

### C. RESULTS

The data were averaged for the 24 *O*s in seven blocks of four trials each for each condition. In Figure 1, errors of omission are plotted; mean latencies are shown in Figure 2.



In the X condition approximately half the signals were auditory and half cutaneous. (Order was determined without restriction from a table of random numbers. Actually, slightly less than half of the signals were cutaneous.) Total errors for the auditory (A), cutaneous (C), simultaneous auditory and cutaneous (AC), and mixed (X) conditions are plotted in Figure 1. The error values shown for the auditory and cutaneous signals presented in the mixed conditions (AX and CX, respectively) are not raw scores; for comparison purposes, the raw scores were multiplied by the reciprocal of their frequency of occurrence in order that the smaller frequency of these signals could be taken into account. (These are hereafter termed adjusted scores.)

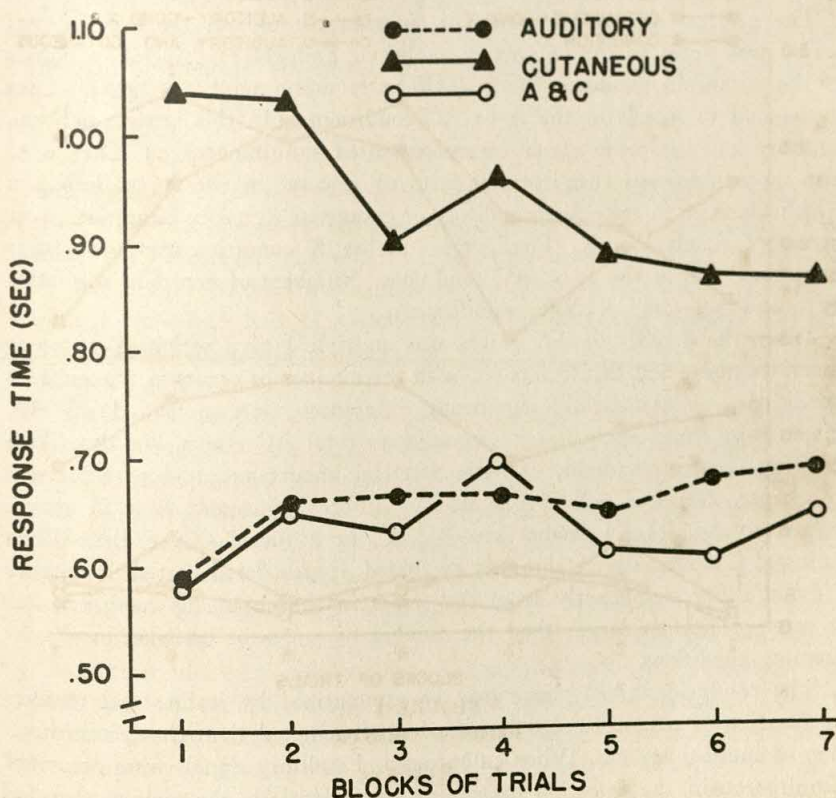


FIGURE 2  
LATENCY ON SUCCESSIVE BLOCKS OF TRIALS

Differences between conditions and between the first and last blocks of four trials were tested by Wilcoxon T tests. Significances of the differences between conditions are listed in Table 1.

TABLE 1  
PROBABILITIES (ESTIMATED BY WILCOXON TS) OF NO TIME DIFFERENCES  
BETWEEN EXPERIMENTAL CONDITIONS

Condition	X	AX	CX	C	A
AC	< .01	< .05	< .01	< .01	NS
X		< .01	< .01	< .01	< .01
AX			< .01	< .01	NS
CX				NS	< .01
C					< .01

The greatest number of errors of omission throughout the experimental period was noted in the cutaneous condition. Failures to respond to signals in the cutaneous condition were significantly more numerous than failures to respond to signals in the A or AC conditions. (In this latter condition, auditory and cutaneous signals were presented simultaneously.) They were also more numerous than total failures to respond in the X condition, in which Os were to respond to auditory or cutaneous signals which were never presented simultaneously. Total errors in the X condition also were more numerous than in the A or AC conditions. Numbers of errors in the latter two conditions did not differ significantly.

When the number of AX errors was multiplied by a weighing factor as described above and then compared with the number of errors in the auditory conditions, no statistically significant differences were noted. Total AX errors were significantly less numerous than total AC errors, but the difference was extremely slight, of little practical importance, and probably due to chance. Adjusted AX errors did not differ significantly from A errors, but were significantly fewer than AC, C, or adjusted CX errors. When adjusted CX errors were similarly evaluated, it was noted that their number did not differ significantly from the errors in the cutaneous condition, but it was significantly larger than the number of errors of omission in the remaining conditions.

The results described above may be summarized by stating that changes in duration of cutaneous signals were less often noted than changes in duration of auditory stimuli. When cutaneous and auditory signals were presented simultaneously, the detection probability was essentially the same as that for auditory signals alone. When they were randomly presented through either channel in the same session, detection probability for each kind of signal was



essentially the same as when the signals were presented alone within a session.

A significant decrease in detection probability over time was noted only for the C condition ( $P < .05$ ). No significant changes were found in the other conditions as a function of time on task. Variability in latency within and between *O*s was considerable. It was apparent from non-parametric tests that the temporal changes observed in this condition did not even approach significance (see Figure 2).

Changes in cutaneous threshold after the sessions are of interest. After the auditory condition, the mean cutaneous threshold significantly decreased (percentage decrease was 24.3,  $P < .01$ ). After the cutaneous and X conditions, the cutaneous threshold increased (percentage increases were 40.0 and 30.6, respectively,  $P < .01$ ). After the AC condition the threshold was lower, but the change was not significant. Changes in auditory matching were not significant.

#### D. DISCUSSION

The findings were not entirely anticipated. Some decline in efficiency in responding to auditory signals had been expected in view of Mackworth's findings (8). However, the relative magnitude of the temporal increments was smaller in Mackworth's experiment, and this may account in large measure for the difference observed. Differences in signal schedules, subject populations, and the extent to which the behavior of observers was monitored in the two experiments may also have produced a differential effect.

It is difficult to explain why there were significantly more errors of omission when monitoring cutaneous signals than when monitoring auditory ones, or why a significant increase in number of errors was observed only for cutaneous signals. It is true that when *O*s were required to detect the *presence of faint* cutaneous or auditory signals, the results were similar; failures of detection were more numerous for cutaneous signals, and a progressive decrement in responding to such signals was noted (6, 7, 10). However, when *O*s were asked to detect the presence of moderate intensity auditory and cutaneous signals, such results were not obtained; detection of both types of signals was efficient and well-maintained. It appears that, though *O*s are able to detect the presence of cutaneous signals and maintain quality of such detection over time, they are not equally able to maintain detection of differences in such signals over time.

The fact that a significant mean change in cutaneous threshold was found after the cutaneous session, a significant decrease after the auditory session,

and no significant change in the mixed modality session is rather provocative. Broadbent's filter theory (1) might predict that as an individual attends to signals in other modalities, and the present results are in accord with this position. The results of the mixed conditions (AC and X) are harder to interpret, but it is uncertain just what would be predicted for such situations. Unfortunately, apparatus limitations at the time of the experiment limited the possibilities of obtaining changes in auditory absolute threshold, but this should be checked in future investigations.

The change in the cutaneous threshold may account for the decrement in responding to cutaneous signals. The differences in trend for detection probability for the auditory and cutaneous probabilities are puzzling. Even if we assume that adaptation or habituation (peripheral or central) is more rapid for cutaneous stimulation, it is not apparent why such processes should vary with modality. Some tentative speculations may be adduced, however.

It was suggested in an earlier article (10) that cutaneous sensations may qualitatively resemble spontaneous (itching, burning, etc.) skin sensations, while the broad-band acoustic noise signals may not resemble "spontaneous" ear noise. The auditory noise apparently is low frequency in quality (2) and would not qualitatively resemble the white noise signals. While the spectrum of cutaneous physiologic noise is not known, cutaneous pitch discrimination is quite poor and is almost impossible to maintain. It is suggested that because of constant exposure to the spontaneous cutaneous noise that central inhibition of excitation resulting from it takes place (9), and that the inhibition may generalize to the cutaneous signals, thus producing the decrement. The auditory noise and signal may be sufficiently different so that such generalization would produce no decrement.

If the intensity of the cutaneous signals is considerable, then presumably the inhibition should not be sufficient to suppress the signal. Indeed the presence of moderately intense cutaneous signals is readily discriminated, and the discrimination is efficiently maintained. It is possible that the subjective intensity is reduced to the point that differential discriminations regarding the signal are impaired, even though detection of the mere presence of the signal is not. It should also be noted that in the present experiment, irrelevant signals were presented quite frequently, and since they differ from the relevant ones only in duration, the habituation process may have been thereby enhanced.

An additional point should be discussed. Buckner and McGrath (4) found that auditory signals were monitored more efficiently than visual ones, just as in this experiment it was found that auditory signals were monitored more efficiently than cutaneous stimuli. However, they found that simultaneous



auditory-visual signals were monitored more efficiently than either auditory or visual signals alone, whereas in the present study the combined auditory-cutaneous signals were monitored no better than auditory signals alone. Moreover, they found that in the mixed condition the detection of the visual signals was enhanced while the detection of the combined signals declined. In the present experiment, the detection of the auditory and cutaneous signals in the mixed session was the same as when the signals were presented unmixed.

The difference may be due to the nature of the task. In their experiment the observers detected an intensity change; in this one, a duration change. Discrepancies, therefore, might be due to differences in difficulty of the discriminations made. It seems probable to the authors that another factor may have been responsible, viz.; the difference in efficiency of monitoring auditory and cutaneous duration changes may be greater than the difference in efficiency for monitoring auditory and visual intensity changes. Perhaps the presence of an inefficiently discriminated stimulus adds very little to the discrimination of a more readily discriminated signal, as the observer may concentrate on the readily discriminated stimulus. Differences in the mixed condition are even harder to interpret. Buckner and McGrath suggested that when two signals differing in detectability are mixed in a session, an intermediate criteria for response is developed. The more detectable signal exceeds the criterion; the less detectable is below it. Therefore, detection of the former is enhanced and that for the latter retarded. Perhaps when detectability of two signals diverges considerably, as in the present experiment, no intermediate criterion for response is developed; rather, two criteria are retained. In this event detection of the signals would not be affected by mixing them.

The latency data need little discussion. Where the observer's task is to respond to a longer signal, he begins to respond as soon as he is reasonably certain that the signal indeed is longer. Under such circumstances, any trend would probably be obscured by intra- and inter-observer variability. The negative finding for changes in latency is therefore of little theoretical significance. For a definitive study of latency change, some discrimination not involving duration would be preferable.

### E. SUMMARY

Twenty-four college students were instructed to attend to auditory, cutaneous, simultaneous auditory and cutaneous, or mixed auditory or cutaneous stimuli presented in separate sessions, and to respond as rapidly as possible by

pressing a key when signals of double length occurred. Detection of auditory signals of this kind was much better than detection of cutaneous signals. Probability of detection for the auditory signals was constant throughout the session, but the probability of detection of cutaneous signals decreased as a function of time on task. Detection of simultaneous auditory and cutaneous signals did not differ appreciably from detection of auditory signals alone. In the session in which the signals could be either auditory or cutaneous, the probability of detection of the auditory and cutaneous signals was approximately the same as their detection probabilities when presented alone. Latency was significantly less for auditory signals than for cutaneous stimuli; but no significant difference in trend in response latency for the two kinds of signals was noted as a function of time on task. A possible interpretation for the results is suggested in terms of differences in physiological noise level, in frequency discrimination, and habituation rate for the auditory and the cutaneous modalities.

## REFERENCES

1. BROADBENT, D. E. *Perception and Communication*. New York: Pergamon, 1958.
2. BROGDEN, W. J., & MILLER, G. A. Physiological noise generated under earphone cushions. *J. Acoust. Soc. Amer.*, 1947, **19**, 620-623.
3. BUCKNER, D. N., HARABEDIAN, A., & McGRATH, J. J. A study of individual differences in individual performance. *Human Factors Res., Inc. Tech Report 2*, Los Angeles, Calif., 1960.
4. BUCKNER, D. N., & McGRATH, J. J. A comparison of performance on single and dual sensory mode vigilance tasks. *Human Factors Res., Inc. Tech Report 8*, Los Angeles, Calif., 1961.
5. ELLIOTT, E. Auditory vigilance tasks. *Adv. Sci.*, 1957, **14**, 353-359.
6. HAWKES, G. R., & LOEB, M. Vigilance for auditory and cutaneous signals. *J. Aud. Res.*, 1961, **1**, 272-284.
7. ———. Vigilance for cutaneous and auditory signals as a function of inter-signal interval and signal strength. *J. Psychol.*, 1962, **53**, 211-218.
8. HAWKES, G. R., BAILEY, R. W., & WARM, J. S. Method and modality in judgments of brief stimulus duration. *J. Aud. Res.*, 1961, **1**, 133-144.
9. LIVINGSTON, R. B. Central control of receptors and sensory transmission systems. In Field, J. and Magoun, H. W. (Eds.), *Handbook of Physiology. Vol. 1*. Baltimore: Williams and Wilkins, 1959, 741-760.
10. LOEB, M., & HAWKES, G. R. Rise and decay time in vigilance for weak auditory and cutaneous stimuli. *Percept. and Mot. Skills*, 1961, **13**, 235-242.
11. MACKWORTH, N. H. *Researches on the Measurement of Human Performance*. London: H. M. Stationery Office, 1950.
12. MARTZ, R. L., & HARRIS, J. D. Signal presentation rate and auditory vigilance. *J. Acoust. Soc. Amer.*, 1961, **33**, 855.
13. SOLANDT, D. Y., & PARTRIDGE, D. M. Research on auditory problems presented by naval operations. *J. Canad. Med. Serv.*, 1946, **3**, 323-329.



14. STEVENS, S. S. The psychophysics of sensory function. In Rosenblith, W. A. (Ed.), *Sensory Communication*. New York: Wiley, 1961, 1-33.

*Psychology Division*

*U. S. Army Medical Research Laboratory*

*Fort Knox, Kentucky*

BINOCULAR PERCEPTUAL DISCRIMINATIONS OF  
AUTHORITY AND PEER GROUP FIGURES  
AMONG OVER, UNDER AND  
EQUAL ACHIEVERS\* .

*Elmira College and Adelphi College*

---

HAL M. WELLS AND DEANNE M. BELL

---

A. INTRODUCTION .

In the last few years, considerable experimentation of binocular perception by stereoscopic and humascope presentation has been conducted. In contrast to earlier research, these experiments have indicated that binocular perception of two different monocular stimuli cannot be attributed strictly to the formal stimulus properties of the two stimuli presented. Hastorf and Myro (3) and Engel (1) demonstrated that a right side up orientation is more frequently perceived than an upside down figure at 1 minute, 0.2 seconds, and 0.1 second exposures. Pettigrew, Allport and Barnett (5) found correlations with the race of the observer and his perception of the race of the stimuli among Afrikaaners. These studies attribute meaningful content, relative to the observer, as the significant variable in the obtained perceptual responses.

In addition to stereoscopic presentation Ittelson has conducted perceptual experiments in the leaf room with aniseikonic glasses and has concluded that the most important person to the observer is least often distorted (4). A preliminary investigation with the *humascope* indicated that in binocular presentation of parental figures at different exposure times perceptual responses included either superimposition or alternation of the two figures, or the exclusive perception of one of them (8).

This present study was designed to answer the following questions: (a) Would the *Os'* responses vary in terms of perception of one or two figures when exposure time remained constant at 0.5 second; (b) Would any one of the three figures, authority, neutral, or peer group, be more frequently perceived than the others; (c) What kind of fusion occurs when one figure is perceived which is not an exclusive perception of a monocular stimulus; and (d) Would there be discriminative differences among over, under, and equal achievers.

---

\* Received in the Editorial Office on April 2, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



Clinical studies have indicated that under and over achievement is related to authority relationships in school, generalized from parental relationships (2, 6, 7). It was hypothesized that discriminative differences would occur between over and equal achievers and under and equal achievers in the perception of authority and peer group figures in the humascope.

## B. METHOD

Each *O* was asked for the name of a friend of his of the same sex whose picture was then taken under controlled conditions with a 35 mm. camera. Photographs were also taken of a man or a woman not known to the *O*s but of the same age who were considered as neutral figures, and two photographs of authority figures, the Dean of Men and the housemother of the girls' dormitory, both of whom serve disciplinary functions.

All the photographs were taken in the same room under comparable conditions and were as alike as possible. The pictures were taken with *Tri X* film and a light meter reading as near *f/8* as possible. They were taken of head and shoulders with body turned slightly to the right. To eliminate facial expression, all the subjects photographed were asked to smile.

Each *O* was shown six stimulus-presentations which included the following pairings: (*a*) neutral and authority; (*b*) authority and neutral; (*c*) friend and neutral; (*d*) neutral and friend; (*e*) authority and friend; and (*f*) friend and authority. The stimulus-presentations included three basic pairings with the left-right order reversed.

Each of the *O*s was shown the six variations in random order, with no two *O*s receiving the same order of presentation. All the photographs shown to the *O*s were of the same sex as the *O*. Neutral and authority pictures of women were shown to the women, and neutral and authority figures of men were shown to the men. The variations in stimulus presentation differed in content only in the photographs of *O*s' friends which were peculiar for each individual.

### 1. Apparatus

The *humascope* was used as the instrument for the binocular perceptual experiment. This machine is an elaboration of the stereoscope. It permits correction for eye dominance by an adjustment device so that points in the two monocular visual fields coincide relative to the individual's vision. Adjustment of the intensity of light in each chamber was used as an additional control. A test pattern consisting of two cards, one with a horizontal black line in the exact center and one with a vertical black line in the exact center, both against

white backgrounds, was used so that each *O* could adjust the angles so they intersected to form a cross. Before each stimulus-presentation, each exposure was flashed on the reverse side of the test pattern cards so the eyes of the photographed persons were exactly on the horizontal lines equidistant from every other side of the card.

## 2. Observers

The *O*s were 33 college sophomores divided into two experimental and one control groups. Eleven of the *O*s were classified as "over achievers," i.e., they were significantly above their predicted rank in their class; 11 of the *O*s were classified as "under achievers," i.e., they were significantly below their predicted rank; and 11 of them, the control group, were classified as "equal achievers," i.e., they fell within the range of their predicted rank. These classifications were determined by a prediction made from their high school rank and the CEEB verbal scores as compared with their actual rank computed upon grade point average. Over numerous years of trying to find an accurate predictive measure, the registrar found that the best predictor consists of a multiple correlation of high school rank and CEEB verbal scores at this college with more than 90 per cent of the students falling within their predicted range. The *O* group, the "over achievers" and the *U* group, the "under achievers," were taken from the extremes of the distribution with a significant statistical deviation ( $p = 0.01$ ).

## 3. Procedure

*O* was seated beneath the binocular viewing device of the humascope. The two card test pattern was placed on the openings opposite *O*, and he was asked to adjust the pattern so the two lines intersected to form a cross. The light remained constant with maximal intensity used in both monocular chambers for each *O*.

The following instructions were given:

You will be shown a series of six exposures. You are to look straight ahead with both eyes and indicate when you are ready. The exposure-time will be short so be prepared to give as complete a description of what you saw as you can. You will be asked to report what you saw after each exposure.

The windows were covered with dark cardboard and the lights turned out except for those in the humascope, and the film projectors. *O*s' responses were tape recorded.

*O* was asked, "What did you just see?" If his reply was vague, e.g., if he



said a man or a woman, he was asked, "Could you describe it a little further?" If a description was given without spontaneous identification, he was asked if he recognized the person. If *O* reported the presence of two figures he was asked if he could elaborate further on their relative positions.

### C. RESULTS

The tape recorded responses were analyzed systematically. The following data were abstracted: (a) each of the six responses was categorized by established criteria; (b) a record was made of the number of times the right or left stimulus was used by *O* in his responses; (c) *O*'s responses were classified by number of times that *O* indicated perception of one or two images; and (d) each of the six stimulus presentations was analyzed separately using the total group's responses.

All responses were categorized in terms of stimulus-recognition which in this case included either identification by name of the stimuli or a correct description of the stimuli. All responses were then classified as (a) Friend, (b) Authority, (c) Neutral, (d) any combination of the afore-mentioned, or (e) Original.

Each of the stimuli was presented four times to each *O* and a total of 44 times to each group. The amount of times that each group correctly recognized the stimulus out of the total amount of times that it appeared was computed for the friend, authority, and neutral figures. Table 1 gives the frequencies and chi square values of the comparisons of stimulus recognition between Group *E* and Group *O*. The values are all statistically significant with Group

TABLE 1  
FREQUENCIES OF STIMULUS RECOGNITION FOR GROUP *E* (EQUAL ACHIEVERS) AND GROUP *O* (OVER ACHIEVERS) AND THE CHI-SQUARE VALUES FOR THE COMPARISON BETWEEN THE TWO GROUPS

Frequencies	Group E	Group O	Chi square values
Recognition of friend	29	39	$\chi^2 = 6.48, p = .02$
Recognition of authority	18	6	$\chi^2 = 8.24, p = .01$
Recognition of neutral	14	2	$\chi^2 = 11.0, p = .01$

*E* giving the most frequent recognition rate of authority and neutral figures and Group *O* giving the most frequent recognition rate of the friend figure. None of the comparisons between Group *B* and Group *E* was significant.

Table 2 indicates the number of responses which included single or double images for Group *E* and Group *O*. Group *E* gave more responses including two images than Group *O*, but more than 70 per cent of the responses of each group described only one image. Table 3 gives the comparison of corresponding double image responses between Group *E* and Group *O* with Group *E* again giving the larger number of corresponding double image responses. The same comparisons between Group *E* and Group *B* were not significant.

TABLE 2  
NUMBER OF SINGLE AND DOUBLE IMAGE RESPONSES FOR GROUP *E*  
AND GROUP *O* WITH CHI SQUARE VALUES

	Single	Double	Chi-square values
Group <i>E</i>	47	19	$\chi^2 = 5.62, p = .02$
Group <i>O</i>	58	8	

TABLE 3  
NUMBER OF CORRESPONDING DOUBLE IMAGE RESPONSES, OR THE CORRECT SIMULTANEOUS  
IDENTIFICATION OF BOTH STIMULI, FOR GROUP *E* AND GROUP *O*  
WITH THE CHI SQUARE VALUE

	Group <i>E</i>	Group <i>O</i>	Chi-square value
Corresponding double images	10	3	$\chi^2 = 4.18, p = .05$

The friend stimulus received the largest total stimulus recognition of 99 times or 75 per cent of the total amount of times that it appeared for the entire group. The authority stimulus was recognized a total of 37 times or 28 per cent of the times it appeared, and the neutral stimulus was only recognized 27 times or 20 per cent of the possible times it could have been correctly identified. The remaining responses were classified either as original, i.e., responses which gave the incorrect identity of someone other than the persons appearing on the photographs or a description of a person whose characteristics did not correspond to those of the stimulus presentation, or there was no indication of the appearance of the figure in *O*'s visual field at all.

These original responses will be considered as perceptual distortions. An analysis of the distortions indicated that they fell into two categories: (a) Peer figures, and (b) Authority figures. Authority distortions occurred only on the four exposures which included the stimulus authority figure and the peer distortion occurred on all six variations. Group *O*, the over achievers, gave the lowest number of corresponding responses and the largest number of original responses, with Group *B* next in order and Group *E* with the lowest number of distortions.



## D. DISCUSSION

The data of this experiment can be classified into the following types of perception: (a) superimposition, the appearance of two separate images; (b) fusion, the appearance of a single image whose description included some of the characteristics of each of the monocular stimuli present; and (c) exclusive perception of a single image. Alternation did not occur in any of *O*s' responses. Less than 30 per cent of the total number of responses for the entire group fell into the superimposition type of response, and only 12 per cent of the total number of responses included correct simultaneous recognition of two separate images.

In spite of the two mechanical controls for eye dominance one might attribute the exclusive perception of one of the two monocular stimuli as a function of eye dominance, except that not one of the 33 *O*s indicated a right or left pattern of responding. Frequently the monocular stimulus perceived was seen with a characteristic of the other stimulus present, e.g., "a picture of my friend with one side of her hair long." The responses which were categorized as original appear to be in part a function of a fusion of the characteristics of both monocular stimuli organized into a stimulus pattern unique to the observer.

The original or distorted responses were categorized into authority and peer group kinds of responses. All the authority distortions, i.e., those appearing in the four variations including pairings of the authority and friend stimuli and the authority and neutral stimuli, which included an adult figure, can be subsumed under authority substitutes. A total of 13 original responses included substitute authority figures including the following: Kefauver, someone's mother, Dave Brubeck, another housemother, a Negro man, and several unidentifiable older women. These responses were given by six different *O*s with eight of the responses occurring within the over-achieving group.

The peer distortions can be classified into three types of distortions: (a) friend identified incorrectly; (b) friend distorted; and (c) peer substitutes. In contrast to the authority distortions, the peer distortions occurred in all six variations even when the friend stimulus was not present. With the exception of one *O*, every *O* recognized their friend at least once during the experiment. Eleven of the *O*s' responses included substitute peer figures which were identified as other students on campus whom the *O*s knew. Eight of the peer distortions occurred when the friend was given as a response when, in fact, the friend stimulus was not presented. The incorrect identification of the friend occurred on the authority and neutral pairings, but the authority figure was never identified when it did not appear. Five of the peer distortions

were classified as distorted views of the friend stimulus, e.g., friend with Jewish nose, friend older, wilted, friend growing evil as in *The Portrait of Dorian Gray*. These five responses were considered as originals inasmuch as they were future interpretations of what the friend might look like. Twenty-three of the responses considered as correct identifications of the friend stimulus did include, however, certain distorted features which appeared to be in part a fusion of the characteristics of both stimuli present. The kind of distortions included the following areas: (a) change in facial expression, e.g., half smiling, stern, unhappy; (b) other apparel, e.g., pearls, glasses, black dress of authority figure; (c) different kinds of hair, e.g., crew cut, gray hair, half blonde; and (d) structural changes, e.g., out of proportion, with three eyes, with two heads.

The data of this experiment indicate that factors other than the formal stimulus properties of the photographs entered into *Os* responses. Meaningful content was not defined in the previous experiments cited; however, Ames has defined it as a kind of weighted average of the past experiences of the perceiver. Meaningful content could be considered in terms of the familiarity of the stimuli to the perceiver which would account for the high recognition rate of the friend stimulus and the low recognition of the neutral stimulus. This interpretation would not hold, however, to the noticeably low recognition rate of the authority stimuli, who were familiar to all the *Os*. It is quite probable that anticipation entered into the perception of the friend stimulus inasmuch as the *Os* knew that their friend was somehow involved in the experiment.

The original responses indicate that the *Os* did organize their percepts into a pattern which had relevance for them. The data indicate that the age of the stimulus figure was a frequently utilized one. The responses which included value descriptions of the stimuli are more clearly a function of the perceiver's orientation. The authority, neutral, and friend stimulus remained constant for each *O*, and yet several of the *Os* perceived their friend in the process of change either in terms of the future, change of mood, or change in physical features. Certain cues from the past experience of the observers were generalized into their discriminations in this setting. It appears that when two monocular cues are different in terms of meaningful content the ambiguity is structured so that the stimulus pattern has meaningful organization to the perceiver.

The question of the substitution of authority and peer figures for the actual authority and peer figures used as stimuli is a difficult problem. In every instance in which this type of response occurred the persons identified in the responses did not correspond in terms of physical features to the stimuli



present. What cues were utilized by the Os which they associated to these particular people cannot be determined in this experiment.

There were no significant differences between the under achievers and equal achievers in perceptual discrimination of authority and peer group figures. The over achievers perceived the peer group figures significantly more often than the equal achievers, and the over achievers perceived the authority figure significantly less frequently than the equal achievers. The over achievers gave the largest number of authority substitutes. Each of the three groups was able to discriminate the peer figure significantly more frequently than the authority and neutral stimulus figures.

### E. SUMMARY

Thirty-three Os, divided into three groups, over achievers, under achievers, and equal achievers, were shown six exposures of authority, peer group, and neutral figures presented binocularly in the humascope. Single image responses were given by each group more frequently than double image responses. The peer figure was recognized more frequently by each group and significantly more frequently by the over-achievers than by the equal achievers. The authority and neutral stimuli were least often recognized by the three groups, but recognized significantly more often by equal achievers than over achievers. Perceptual distortion occurred which included substitute authority and peer group figures and various distortions of the friend stimulus. The over achievers gave the largest number of distorted responses.

### REFERENCES

1. ENGEL, E. The role of content in binocular resolution. *Amer. J. Psychol.*, 1956, **69**, 87-91.
2. FLIEGLER, L. A. Understanding the underachieving gifted child. *Psychol. Rep.*, 1957, **3**, 533-536.
3. HASTORF, A. H., & MYRO, G. The effect of meaning on binocular rivalry. *Amer. J. Psychol.*, 1959, **72**, 395-400.
4. ITTELSON, W. H. *Visual Perception*, 1960.
5. PETTIGREW, T. F., ALLPORT, G. W., & BARNETT, E. D. Binocular resolution and perception of race in South Africa. *Brit. J. Psychol.*, 1958, **49**, 265-278.
6. SPERRY, B., STOVER, N., REINER, B., & ULRICH, D. Renunciation and denial in learning difficulties. *Amer. J. Orthopsych.*, 1958, **28**, 98-111.
7. SUZIEDELIS, A. Some adjustment correlates of the self and the ideal self concepts as influenced by perceived parental values. Washington: Catholic Univ. Press, 1958.
8. WELLS, H. M., & BELL, D. M. Binocular rivalry of parental figures in the humascope. *J. Psychol.*, 1960, **50**, 145-152.

*Psychological Laboratory*  
*Elmira College*  
*Elmira, New York*

## VERBAL REPORT AND JUDGMENT OF AN UNSTRUCTURED STIMULUS\*

*Institute of Group Relations, The University of Oklahoma*

---

WILLIAM ROBERT HOOD AND MUZAHER SHERIF

---

### A. INTRODUCTION AND PROBLEM

One of the established experimental methods in social psychology is measurement of changes in judgment accompanying variations in social influence. Procedures have been applied to judgment of a variety of stimulus dimensions (e.g., apparent movement, warmth, length, numerosity, affectivity) and have utilized different social influences (e.g., judgments of other persons in various prestige and status relationships to *S*, or with membership in social aggregates and groups).

During the last decade, interest in psychological processes involved in behavioral changes and a healthy concern with the validity of laboratory results have led to certain pertinent questions. Does the individual who changes his verbal reports with the introduction of a social influence actually *see* the stimulus situation differently, or is he simply reporting in a way calculated to avoid disapproval and to appear agreeable? Is there a discrepancy between his perception or judgment of the situation and his verbal report to *E*? That discrepancies do occur is shown by studies in which stimuli to be judged clearly differ and *S* faces a number of other individuals (planted *Ss*) stating, in effect, that they do not differ (1, pp. 451-483.) Here the typical finding is that a minority of the *Ss* agree with the planted *Ss*' judgments and that most of these later say that they did not in fact "see" the situation as they reported it.

Festinger (3) suggested a behavioral criterion for evaluating whether or not a discrepancy between *S*'s experience and his verbal report occurs, viz., determining whether or not convergence to an experimentally introduced standard continues following its removal. On the basis of this criterion, experimental evidence of a discrepancy is lacking when the stimuli are unstructured in the dimension judged, i.e., when stimulus determinants are such that alternative perceptual organizations are feasible. The original studies on judgments of extent of autokinetic movement in interaction situa-

---

\* Received in the Editorial Office on April 2, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



tions (6) revealed no discrepancy between judgment and verbal report by this criterion. After judging extent of movement with others, *Ss* subsequently adhered to the range and norm developed in the interaction situation when making judgments alone on a different day. More recently, the maintenance of judgments formed in social situations was reported in individual sessions held after intervals of 28 days (2) and a year (5).

If lack of correspondence between judgment and verbal report were the general finding in experiments on social influence, it might be concluded that such experiments reveal merely responses expected of agreeable, socialized individuals to direct social pressures or demands from others in the situation. If so, findings obtained by this laboratory method would not be pertinent to the more lasting changes produced by social influences in actual life. Thus the correspondence or lack of correspondence between judgment and verbal report is a crucial problem in social psychology. The problem may be stated more specifically: Under what conditions is a discrepancy found or not found? What is the nature of the social influence in these conditions and how is it perceived by *S*?

As a first step, it may be noted that discrepancies are sometimes reported in situations where the stimulus dimension is definitely graded in clear alternatives (structured situations). Judgments of other individuals contrary to perceived stimulus differences contribute to *S*'s perception of "social pressures" to conform. At an opposite pole, when facing a situation whose stimulus determinants permit many alternatives, *S* is influenced by an experimentally introduced standard and may be unaware of the influence. In the latter situation, it may be misleading to speak of "social pressures." The problem of the present experiment concerns this latter pole, which may be conceived as an end segment in the range of gradations of structure in social influence situations.

The experimental situation was defined by the following guides: (a) The stimulus dimension judged permits varied alternative modes of experience and behavior. (The autokinetic situation was chosen.) (b) Procedures eliminate as much as possible any suspicion that the experiment has anything to do with any kind of social influence. (c) Immediate social pressures, in the form of the presence of another person making judgments or the sound of his voice, are eliminated. The social influence is perceived by *S* as coincidental to his presence in the situation and is absent when he renders judgments. (d) *S* does not "commit" himself in the presence of a planted *S* and is encouraged to judge as he sees fit. (e) Extensive data are collected on *S*'s experiences in the experiment as reported after he makes judgments. Such

data are a basis, in fact the only available basis, for inferring his experience at the time.

The hypotheses were that under the conditions stated above: (a) *S*'s judgments would converge toward those of a planted *S* to which he was exposed prior to rendering judgment, and (b) there would be no discrepancy between experience of the stimulus and verbally reported judgments.

## B. METHOD

*Ss* were told that the purpose of the experiment was the design of tests of visual abilities under low illumination. *E* identified himself as a graduate student employed to work on a research contract and avoided any mention of psychology. The experimental darkroom was located in an Army Reserve building. Before entering the darkroom, *E* explained to *S*:

We've found it takes the average person's eyes about three to 15 minutes to adjust to the dark enough to see the light. I left the observer who started before you in the darkroom while I came to get you so that we could save that 15 minutes. He has about that many estimates to make in order to finish his series. I'll bet he's asleep by now. You can just sit in the darkroom and let your eyes adapt while he finishes, and then you and I will be ready to go to work.

The room was dark when *S* entered and he never saw the planted *S*. *E* asked the planted *S* if he had taken a nap. The planted *S* answered: "No, but this would be a good place to take one. It sure is dark in here." Except for his judgments, he said nothing else in the naive *S*'s presence. The naive *S* was led to a chair at one side of a table where the planted *S* was seated, facing the light source. Then he was told:

I'll tell you what we are doing before we start again so that it won't be quite so boring—just sitting there. The observer who is working now has a box in front of him with a button on it. His job is to watch for the light to come on down at the other end of the room—off to your right. I will say "ready" about three seconds before the light appears. He watches the light, and as soon as it starts to move, he punches the button. After a few seconds the light goes off, and then he tells me how far it moved—just the distance it moved through space; the direction doesn't matter.

The planted *S* then gave a series of 18 judgments, spoken in a clear voice but with a degree of assurance planned to approximate that conveyed by naive *Ss* in pre-tests. For half of the *Ss*, these 18 judgments were distributed between 1" and 5" with a mode of 3". For the other half of the *Ss*, the 18 judgments were distributed between 6" and 10" with a mode of 8". When



the series of 18 judgments was completed, *S* was asked to wait in the dark-room (two minutes), while the other observer was paid. Then *S* moved behind the table and the instructions for judging movement were repeated. *E* avoided any comment on the planted *S*'s judgments. A few *S*s expressed disagreement with them after the "plant" left, and they were told: "You call them the way *you* see them."

Following four trial judgments, each *S* gave a series of 50 judgments at one minute intervals. Then *S* was taken to where a questionnaire was administered.

### C. APPARATUS

The autokinetic apparatus exposed a pinpoint of light automatically through a circular hole 1 mm. in diameter (see Sherif and Harvey, 1952, for specifications). The brilliance of the light was reduced approximately to half, and exposure time was set at two seconds. Pretests of these adjustments resulted in relatively small and consistent judgments of extent of apparent movement. The darkroom was 15'  $\times$  30' with *S*-to-light distance of 20'.

The questionnaire administered following the judgment included several linear scales to secure estimates of confidence, influence of the planted *S* and extent of agreement with him, as well as open-ended items concerning purpose and possible utility of the experiment, role of the planted *S*, descriptions and estimates of apparent movement.

### D. SUBJECTS

*S*s were 24 male college students paid as "observers" and unfamiliar with the autokinetic phenomenon. Random assignment to the two experimental conditions was restricted only by matching for age. Pre-test data indicated differences in response to the experimental conditions related to age, which factor probably summarizes several related characteristics of college students, e.g., classification, academic success, ease in an experimental situation. In each condition, five *S*s were between 17 and 19 years old, two *S*s were 20 years, and five *S*s between 21 and 27 years. The planted *S* was the same person in each session (age, 20 years).

### E. RESULTS

Data bearing on three general questions are pertinent to the hypotheses:

1. Do *S*'s judgments conform to the mode and range of the planted *S*'s judgments previously overheard?
2. Do *S*'s experiences as reported immediately after the experimental

session agree with his verbally reported judgments?

3. Is *S* aware of any influence from the prior judgments of the planted *S*?

Table 1 gives the median judgments by *S*s in each condition. The median of all judgments for the sample exposed to the 1-5" range is 3.98", and that of the sample exposed to the 6-10" range is 6.79".

TABLE 1  
MEDIAN JUDGMENTS BY *S*s INITIALLY EXPOSED TO DIFFERENT RANGES

Condition: 1-5" Range Inches	Condition: 6-10" Range Inches
2.6	3.1
2.7	4.3
3.2	4.9
3.4	6.5
3.5	6.6
3.8	6.8
3.9	6.8
4.1	7.5
4.3	7.6
4.9	7.7
5.1	8.5
5.7	11.5

Mann-Whitney  $U = 16$   
 $p < .001$  (one-tailed test)

The differences between the medians for the two conditions were tested by the Mann-Whitney statistic,  $U$ , and were significant ( $p < .001$ ).

Figure 1 shows the proportion of *S*s' judgments which were 5" or less and 6" or more. *S*s exposed to the 1-5" range placed the bulk of their judgments below 5" and *S*s exposed to the 6-10" range concentrated their judgments above 6". In the 1-5" condition, 81.1 per cent of the judgments fell between 1-5" and the range of judgments was 10". In the second condition, 70.5 per cent of judgments were 6" or greater, and the range was 24".

Responses to questionnaire items concerning the purpose of the experiment were identical or entirely consistent with the instructions for all but one *S*. The latter, a bright young physics student, had already observed to *E* that overhearing another person might affect one's judgments. However, this *S* estimated that he himself was influenced by the plant in only about 25 per cent of his judgments. Actually 74 per cent of his judgments fell within the plant's range. Explanations of the procedures by other *S*s were in terms of tests for "night flying," "night driving" and the like. All *S*s accepted *E*'s explanation of the plant's presence.

In inferring correspondence between *S*s' verbal reports and experiences of



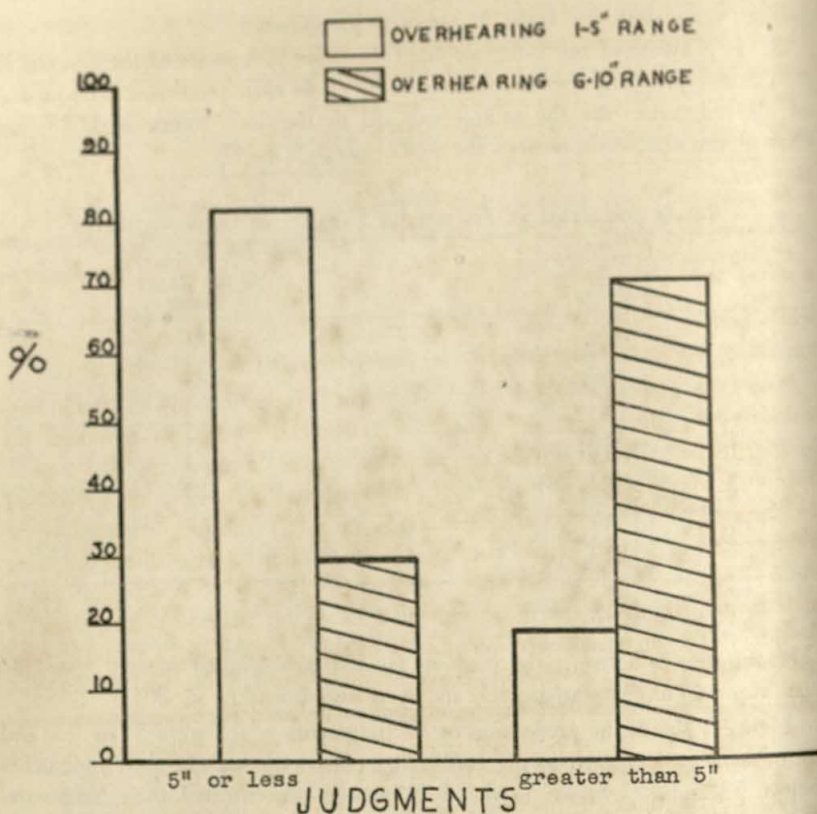


FIGURE 1

PROPORTIONS OF JUDGMENTS 5" OR LESS AND GREATER THAN 5" FOR TWO CONDITIONS

apparent movement, a crucial comparison is between medians of the obtained judgments and *Ss'* responses to the questionnaire item: How far did the light seem to move usually?" The summary in Table 2 reveals small differences between the two. The significance of differences between each *S's* median judgment and his subsequent report of "usual" extent of perceived movement was tested by the Wilcoxon Matched-Pairs Signed-Ranks Test. The resulting *Ts* of 21 and 26 for *Ss* in the 1-5" and 6-10" conditions, respectively, were not significant ( $p > .05$ , two-tailed test). Thus the hypothesis that *Ss* reported movement as they saw it is supported.

*Ss* accepted the planted *S*, whom they never saw, as another student like themselves. Only one of the 24 *Ss* believed that the plant judged accurately,

TABLE 2  
 MEDIAN OF OBTAINED JUDGMENTS AND EXTENT OF MOVEMENT  
 USUALLY PERCEIVED BY *S*s

Condition	Obtained judgments Inches	Usual extent perceived Inches	T*	Range of differences** Inches
1-5" range	3.98	3.88	21	-1.2-+3.3
6-10" range	6.79	7.00	26	-1.5-+1.3

\* Wilcoxon test:  $p > .05$ .

\*\* Difference between *S*'s median judgment and usual extent of movement he reported having perceived.

while another stated that he was uncertain as to his accuracy. The remaining 22 *S*s stated either that the plant overestimated or underestimated.

Asked to estimate on a linear scale the proportion of his judgments influenced by the plant, only one *S* overestimated the extent to which his judgments actually fell within the prescribed range of the plant. The median difference between the proportion of judgments falling within the prescribed 1-5" range and the proportion which *S*s in that condition estimated to be influenced was 64 per cent (range: -7-100). The median discrepancy for the 6-10" condition was 57.5 per cent (range: 16-80). Since all but one *S* underestimated the extent of agreement between his judgments and those of the plant, the direction of differences is clearly significant ( $p < .005$ , two-tailed Wilcoxon Matched-Pairs Signed-Ranks Test). The size of the discrepancy between the proportions of judgments actually falling within the prescribed ranges and the proportions *S*s thought were influenced was significant for both conditions ( $p = .005$ , two-tailed Walsh test).

The distributions of responses on a 5-category scale concerning amount of agreement between plant's judgments and *S*'s own judgments were practically identical for the two conditions. One *S* in each condition rated his agreement with the plant between "Agreed some of the time" and "Much agreement." Of the remaining 22 *S*s, half checked "Agreed some of the time" and half checked "Positively no agreement" or "Little agreement."

#### F. DISCUSSION

Analysis of the results has indicated that the prescribed range of a planted *S*, overheard briefly prior to rendering judgments, did influence the norm and range adopted by *S* in this situation. The purpose of the experiment and the role of the plant were accepted as genuine. There was no significant discrepancy between *S*'s reported judgments and his experience of extent of movement. On the whole, *S*s were not aware of the extent to which their



judgments coincided with the plant's. Thus we may consider the main hypotheses supported. In a situation lacking clearcut determinants in the stimulus dimension judged, the spoken judgments of another person serve to anchor experience of the individual, and this experience is accurately reflected in his verbal reports while rendering judgment.

While much experimentation on social influence has quite properly focussed on interaction situations, it is apparent that even in the absence of interaction between individuals or the immediate sound of another's voice, the judgments of another person overheard previously may exert a categorizing effect, delimiting the scope of perceived alternatives and affecting the modal value. This conclusion is consistent with the experimental findings concerning the categorizing effect of linguistic concepts (4). In such situations, it would seem misleading to conceptualize the influence process in terms of arbitrary "social pressures," particularly when *S* is unaware of being influenced.

The situation in this study represents near-minimal social influence whose strengthening by the ascription of authority or prestige should lead to even more striking differences in judgments by *Ss* overhearing different prescribed ranges.

The use of a minimal social influence in the present experiment permits analysis of some of the stimulus determinants operating even in so unstructured a situation as the autokinetic set-up. Under the usual conditions of autokinetic experiments as to size of room, distance from light, length of exposure and brilliance of light, *Ss* ordinarily center judgments around a mode of 3-6", the individual who exceeds a mode of 8-10" being exceptional. In comparison, judgments made in a large, empty auditorium are greater in both mode and range (7). In the present experiment, the brilliance of the light was reduced, which resulted in pre-test judgments of small movement within a narrow range.

Certain differences in the responses by *Ss* exposed to the 1-5" range and the 6-10" range are worth noting. In a situation more conducive to the perception of small movements, the introduction of the 6-10" range not only set a level considerably higher than "normal" for these viewing conditions, but also resulted in a range (24") unusually large for these conditions. In the autokinetic situation or other situations conducive to various alternative modes of perceptual patterning, possible conflict among alternatives can be resolved rather easily. In the present case, *Ss* in the 6-10" condition estimated distance to the light as farther (median: 19') than *Ss* in the 1-5" condition (median: 14'). This estimate is congruent with their larger median judgment of perceived movement (*viz.*, 6.79", as compared to 3.98"). The

limits imposed by stimulus determinants are strikingly apparent when the social influence diverges from them increasingly. Whittaker (8) has shown that when the social influence diverges widely from the *S*'s mode, he ignores the social standard completely.

### G. SUMMARY

The validity of experimental methods for studying lasting changes in judgment as a function of social influence requires that verbal reports of judgment reflect *S*'s experience of the stimulus dimension. This paper discusses factors determining whether or not a discrepancy between verbal reports and *S*'s experience will arise. An experiment is presented in which the stimulus dimension judged (extent of autokinetic movement) permitted alternative modes of experience and report. Procedures were designed so that *S*s would not suspect that a study of social influence was in progress. *S* did not "commit" himself in the presence of another person, and immediate "social pressures" were eliminated. *S* was exposed, as though accidentally, to the prescribed judgments of a planted *S* prior to rendering judgment.

In this experimental situation, judgments of *S*s exposed initially to different prescribed ranges did differ significantly in the predicted directions. There was no significant difference between verbal reports in rendering judgment and experiences of movement, as inferred from other behavioral data. On the whole, *S*s accepted *E*'s account of the purpose and procedures as genuine, and were unaware of the extent to which their judgments had been influenced.

This situation is contrasted with those at an opposite pole, in which determinants in the physical stimulus situation are predominant and discrepant social standards are introduced.

### REFERENCES

1. ASCH, S. E. *Social Psychology*. New York: Prentice-Hall, 1952.
2. BOVARD, E. W., Jr. Social norms and the individual. *J. Abnorm. & Soc. Psychol.*, 1948, **43**, 62-69.
3. FESTINGER, L. An analysis of complaint behavior. In Sherif, M. & Wilson, M. O. (Eds.), *Group Relations at the Crossroads*. New York: Harper, 1953, 232-256.
4. GIBSON, J. J. The reproduction of visually perceived forms. *J. Exp. Psychol.*, 1929, **12**, 127-155.
5. ROHRER, J. H., BARON, S. H., HOFFMAN, E. L., & SWANDER, D. V. The stability of autokinetic judgments. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 595-597.
6. SHERIF, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, No. 187.
7. SHERIF, M., & HARVEY, O. J. A study in ego functioning: Elimination of stable anchorages in individual and group situations. *Sociometry*, 1952, **15**, 272-305.



8. WHITTAKER, J. O. The effects of experimentally introduced anchorages upon judgments in the autokinetic situation. Unpublished doctoral dissertation, University of Oklahoma, 1958.

*Institute of Group Relations  
The University of Oklahoma  
Norman, Oklahoma*

## A COMPARATIVE STUDY OF THE EFFECTS OF LISTICA AND MEPROBAMATE UPON MOTOR FUNCTIONING\*

Collegeville, Pennsylvania

M. POWELL LAWTON AND BURTON CAHN

### A. PURPOSE

Hydroxyphenamate (Listica)<sup>1</sup> is a new tranquilizing agent for mild and moderate anxiety states which are met with in the office practice of medicine. Hydroxyphenamate is a monocarbamate, in distinction to meprobamate, which is a dicarbamate. The effect of the drug on the central nervous system is caused by the molecule as a whole, rather than by the carbamate group alone. There is no evidence to indicate that a tolerance or cumulative effect is developed. It is not as yet known whether the influence of hydroxyphenamate is directly upon the cortex or whether it is through the indirect action of subcortical areas which in turn affect the cortex. Clinical improvement has been reported with hydroxyphenamate in anxiety states (1), dermatologic cases (2), allergies (3), cardiovascular illness (6), headache (4), and alcoholism (5). The present study compared the effects of hydroxyphenamate and meprobamate upon psychomotor functions.

### B. METHOD

The following measures of motor functioning were used:

1. Cancellation. The subject (*S*) was told to cross out as many "e's" on a printed page as he could in 60 minutes. This task was repeated, and a single score obtained, as well as an omission (error) score.

2. Digit Symbol. *S* had a code key of symbols paired with digits 1 through 9. He was asked to fill in as many of 90 empty squares as he could in 90 minutes, referring to the key as much as he wished. This is a subtest of the Wechsler Scales.

3. Dotting Test. *S* was presented with an  $8 \times 25$  grid of quarter-inch squares and was asked to put a single dot in as many squares as he could in 60 minutes. This added an element of motor control to the standard speed of tapping test. The test was repeated and a single score obtained.

\* Received in the Editorial Office on April 9, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Listica was supplied by the Armour Pharmaceutical Corporation, Kankakee, Illinois.



4. Purdue Pegboard. This is a standardized test which requires *S* to work as quickly as possible placing by hand small metal pegs in holes.

a. Right hand speed. *S* placed as many pegs as possible in 30 seconds. The test was done three times and a single score obtained.

b. Both hands assembly. *S* placed a peg in hole, a washer on the peg, a collar on the washer, and another washer on top of the collar, assembling as many as possible in 60 seconds. The test was done three times, and a single score obtained.

5. Porteus Mazes. This was a series of three paper and pencil mazes graded for difficulty and scored for time and errors.

6. Steadiness Test. The *S* attempted to hold a metal rod inside three holes of varying diameters without touching the edges of the metal plate. A buzzer signaled each contact and the score was number of contacts.

These tests represent the simplest motor behavior (dotting, cancellation and right hand speed), moderately complex behavior (both hands assembly), and complex motor behavior which has many central aspects of choice and discrimination (digit symbol and mazes). The ability measured by the steadiness test is probably not on the same continuum, though steadiness and relaxation with moderate tonus are necessary components of motor coordination. The complex tasks, digit symbol and mazes, bear a relatively high relationship to general intelligence. Most of these tests have been used repeatedly in drug evaluations.

*Ss* used were 48 volunteers from a city prison, divided randomly into three groups of sixteen each. These groups did not differ significantly in mean years of school completed (9.9, 8.9 and 9.4), percentages of rurally educated (about 12 per cent of each group), or mean age (25.4, 28.7 and 27.9).

The design of the study called first for all tests to be administered to each group on the day preceding treatment (Day 0). The same tests were repeated on one group (Listica-meprobamate) after four days of treatment with Listica, on another group after four days of treatment with meprobamate (meprobamate-Listica) and on a third group after four days of medication with identical placebo pills (Day 4). After testing, the medication was reversed, and the Listica-meprobamate group was tested following four days of treatment with meprobamate, the meprobamate-Listica group was tested following four days treatment with Listica and the placebo group again retested following eight days of medication with placebo (Day 8). Three *Ss* in the meprobamate-Listica group were discharged after being tested on Day 4. Therefore this group was reduced to a size of thirteen for several of the analyses.

The author was not aware of the experimental construction of the groups until testing was completed.

The design thus allowed for the following analyses:

1. Assuming that the three groups were equal before the study (that is, on Day 0), did either experimental group differ from the control group on Day 4? Or alternately, was either active drug associated with a change from Day 0 to Day 4 greater than that associated with medication by placebo? Was there such an effect when the two drug groups were compared to the control from Day 0 to Day 8?

2. Did the two active drugs differ from each other in their effects upon psychomotor coordination? Did the Listica group differ from the mepro-bamate group on Day 4? Was there any significant reversal of trend which took place when the drugs were switched, that is, from Day 4 to Day 8?

The final choice of indices of motor performance was in part determined by the empirical reliabilities and distributions of the measures. On the cancellation task, the distribution of number correct was quite different from the distribution of errors. No combination of number right minus errors gave as wide a distribution or as high a test-retest reliability as did each taken alone. Therefore, they may be presumed to measure different skills and were both used in the analysis.

The mazes were also scored for total time, number of errors, and several combinations of time and errors. Again, both time and error scores were retained, as they seemed relatively independent statistically.

TABLE 1  
TEST RELIABILITIES (PLACEBO GROUP, DAY 0-DAY 4)

Test	Rank order correlation
Cancellation	.71
Digit symbol	.78
Dotting	.89
Right hand	.75
Assembly	.66
Cancellation errors	.71
Steadiness	.84
Maze time	.74
Maze errors	.68

The test-retest reliabilities, derived from the sixteen Ss in the placebo group, are shown in Table 1. They appear to be satisfactory for statistical analysis.

Some of the scores were normally distributed, but the data for the cancellation errors, steadiness, maze time and maze errors were markedly



skewed, so that nonparametric statistics had to be employed. The .10 level of significance was used as one which would be best suited to an exploratory investigation, using two-tailed tests in every instance, since no a priori hypotheses were made regarding which group would be superior.

### C. RESULTS

Table 2 indicates that no significant premedication differences existed among groups, using the analysis of variance ( $F$ ) for the normally distributed variables and the Kruskal-Wallis one-way analysis of variance technique ( $H$ ) for the skewed variables. Probably values ( $p$ ) are indicated in this and all other tables only if they are less than the .10 level.

Table 3 lists mean and median scores after four days of medication. Again, no significant differences appeared, suggesting that both controls and medicated groups improved approximately equally on their second tests. Table 4

TABLE 2  
DAY 0 SCORES

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	85.1	82.4	79.2	.43
Digit symbol	36.1	36.6	37.9	.20
Dotting	229.7	225.8	234.4	.15
Right hand	45.3	44.3	47.3	1.92
Assembly	107.2	102.8	111.9	1.07
		<i>Medians</i>		<i>H</i>
Cancellation Errors	15.0	19.0	9.5	1.5
Steadiness	34.5	31.0	30.0	.7
Maze time	351.5	248.0	244.0	1.5
Maze errors	12.0	7.0	6.0	3.5

TABLE 3  
DAY 4 SCORES

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	86.1	86.7	84.4	.07
Digit symbol	42.1	39.3	41.4	.45
Dotting	247.4	237.7	242.8	.18
Right hand	47.2	46.8	46.9	1.1
Assembly	119.1	114.7	118.4	.27
		<i>Medians</i>		<i>H</i>
Cancellation errors	7.5	13.5	11.0	1.6
Steadiness	25.0	25.0	19.0	2.9
Maze time	178.5	239.5	224.0	0.4
Maze errors	10.0	9.5	8.5	0.7

shows mean or median changes from Day 0 to Day 4, arranged in such a manner that a positive score indicates improvement and a negative score a decrement. One notes that the placebo group improved significantly more than either experimental group in total time required to trace the three mazes. Thus, while the three groups do not differ in absolute level of maze performance on Day 4, the trend toward a longer total time by the control group on Day 0 was reversed by their greater improvement.

Table 5 shows mean and median scores for Day 8, after the control group

TABLE 4  
DIFFERENCE SCORES DAY 0 TO DAY 4

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	1.00	4.25	5.25	1.06
Digit symbol	5.94	2.75	3.38	2.59
Dotting	17.75	12.00	8.38	.74
Right hand	1.94	2.44	.42	2.03
Assembly	12.3	11.9	6.4	1.30
		<i>Medians</i>		<i>H p</i>
Cancellation errors	5.0	3.0	2.0	2.3
Steadiness	6.0	16.0	15.5	3.6
Maze time	82.0	14.5	4.5	10.1 < .01
Maze errors	1.5	0.5	1.5	1.8

TABLE 5  
DAY 8 SCORES

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	88.1	90.2	87.2	.12
Digit symbol	45.1	44.3	45.7	.08
Dotting	247.3	249.2	248.2	.01
Right hand	49.0	48.5	49.1	.06
Assembly	125.3	119.2	122.6	.38
		<i>Medians</i>		<i>H p</i>
Cancellation errors	7.0	7.0	11.0	1.6
Steadiness	19.5	9.0	9.0	5.2 < .10
Maze time	158.5	185.0	157.5	0.1
Maze errors	7.5	3.0	6.5	5.7 < .10

had taken placebos for seven days, and the original Listica and meprobamate groups had switched medications. On Day 8 there were marginally significant differences among groups in steadiness and maze errors. The two medicated groups were superior to the control group in hand steadiness, while the group currently on Listica committed fewer maze errors than did



either the control or the meprobamate group. These two changes were apparently cumulative from Day 0, however, since Table 6 indicates that no significant change occurred from Day 4 to Day 8.

TABLE 6  
DIFFERENCE SCORES: DAY 4-DAY 8

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	2.00	4.31	2.81	.43
Digit symbol	3.06	4.14	4.25	.46
Dotting	—13	14.00	4.25	3.01
Right hand	1.81	1.08	—81	.26
Assembly	6.19	3.31	4.13	.33
		<i>Medians</i>		
Cancellation errors	4.0	2.0	0.5	.7
Steadiness	10.5	10.0	5.5	.2
Maze time	45.0	52.0	62.0	1.1
Maze errors	5.0	2.0	0.5	3.1

TABLE 7  
DIFFERENCE SCORES: DAY 0-DAY 8

	Placebo	Meprobamate- Listica	Listica- Meprobamate	
		<i>Means</i>		<i>F</i>
Cancellation	3.00	8.54	8.00	2.23
Digit symbol	9.00	7.38	7.88	.35
Dotting	17.60	26.62	13.69	1.11
Right hand	3.75	4.00	1.75	2.11
Assembly	18.06	16.92	10.69	2.95
		<i>Medians</i>		
Cancellation errors	8.0	5.92	3.5	0.9 <i>p</i>
Steadiness	16.0	26.0	23.0	4.8 < .10
Maze time	134.0	109.0	58.0	3.7
Maze errors	5.5	2.0	0.5	5.3 < .10

Table 7, showing differences from Day 0 to Day 8, is of unclear meaning in comparing the Listica group with the meprobamate group. The two differences of marginal statistical significance show, however, that the two experimental groups did improve more in steadiness than did the control, but that the control group improved more in number of maze errors.

#### D. DISCUSSION

The occurrence of five significant differences at the .10 level out of a total of fifty-four tests performed is not impressive and lengthy discussion of the meaning of the differences does not seem indicated. The fact that the

control group more than made up for a tendency toward an initial deficit in maze performance (both in time and errors) could mean either that the initial deficit was a chance phenomenon or that both meprobamate and Listica impeded improvement on this task. The fact that no other test reflected a similar tendency argues against the latter interpretation of this isolated finding.

The greater improvement in steadiness, as seen in final scores (Day 8) and improvement scores (Day 0-Day 8 differences), is statistically marginal, and may therefore be a chance variation. It is of interest, however, that these differences are consistent with the purported effects of the medication, that is, muscular relaxation.

### E. CONCLUSIONS

Through most of the range of muscular coordination which was investigated, there is no pronouncedly deleterious effect of either Listica or meprobamate. Simple and complex manual tasks and muscular performance tasks with some central components were used, with only one of the latter tasks showing any suspicion of decrement. There was some possibility of improvement indicated by the superiority of both experimental groups in the steadiness task.

Therefore, the present study suggests that hand muscle coordination remains efficient with the administration of Listica, as well as with meprobamate.

### REFERENCES

1. CAHN, B. Effect of hydroxyphenamate in the treatment of mild and moderate anxiety states. *Dis. Nerv. Syst.*, 1961, **22**, 30-32 (Supplement).
2. CAHN, M. M., & LEVY, E. J. Use of hydroxyphenamate (Listica) in dermatologic therapy. *Dis. Nerv. Syst.* 1961, **22**, 47-49 (Supplement).
3. EISENBERG, B. Amelioration of allergic symptoms with a new tranquillizing drug (Listica). *Dis. Nerv. Syst.*, 1961, **22**, 25-29 (Supplement).
4. FRIEDMAN, A. Pharmacological approach to the treatment of headache. *Dis. Nerv. Syst.*, 1961, **22**, 36-40 (Supplement).
5. GOULDMAN, C., LUNDE, F., & DAVIS, J. Clinical trial of hydroxyphenamate in alcoholic patients. *Dis. Nerv. Syst.*, 1961, **22**, 44-46 (Supplement).
6. GREENSPAN, E. Use of hydroxyphenamate in some forms of cardiovascular disease. *Dis. Nerv. Syst.*, 1961, **22**, 33-35 (Supplement).

R.D. #2

Collegeville, Pennsylvania



## RACIAL FRICTION IN THE DEEP SOUTH\*<sup>1</sup>

*Department of Psychology, The University of Mississippi*

JOHN B. WOLFE AND PAUL HORN

### A. INTRODUCTION

Burrowing owls, rattlesnakes, and prairie dogs, according to certain storytellers, live happily together in the same burrow even though the owls and snakes feed on the baby dogs. We were reminded of this anecdote by some of the results of a study we made recently of Negroes and whites living in the same village in the deep south. This analogy was surprising since the study had been designed to reveal causes of inter-racial unhappiness and since inter-racial unhappiness at the level of individual face-to-face conduct had been assumed. The assumption was based on our exposure over the years to disguised and undisguised psychoanalytic theory, *Frustration and Aggression*, culture analysis, attitude studies, and the press. Thus Davis, Gardner, and Gardner (1) state that southern Negroes are required to show enthusiastic compliance in dealings with the whites if they are to avoid grief. Hortense Powdermaker (2) suggests that the Negro must, in effect, behave as a masochist if he is to get along. The *New York Times*, March 8, 1958, quotes William Faulkner as saying, "It's always been my belief that white folks and colored folks simply don't like one another."

The purpose of our study was to identify causes of strained racial relations by ascertaining *specific incidents of friction* that occurred in the lives of individual Negroes in their contacts with whites, and vice versa. The incidents were revealed in quasi-depth interviews controlled by interviewers who used memorized questions after rapport had been established. The main questions sought precise descriptions of disturbing situations that had arisen between the interviewee and a member of the opposite race. The interviewer secured from each respondent detailed accounts of two incidents that were most unpleasant and, for another purpose, two incidents that were most pleasant. As control measures for the interpretation of the data, incidents of both displeasure and pleasure involving members of the respondent's own race were also obtained. An open question terminated the interview. It was

\* Received in the Editorial Office on April 9, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Financial support for this research was furnished by The Human Genetics Fund of New York City and was disbursed through The University of Mississippi.

"What do you really think is the most objectionable thing about the *opposite* race?"

The study, viewed practically, sought a list of trouble-causing incidents ranked in order of frequency of occurrence. It was hoped that the list in the hands or "minds" of culture carriers, might be a useful guide.

### B. SUBJECTS

A total of 1196 persons were interviewed. They were, in the main, native to the mid-south and were at home or at college when interviewed. Of the total, 540 were Negroes; 656 were whites. The Negroes were 136 college men, 89 college women, 150 rural men, and 165 rural women. The whites were 180 college men, 151 college women, 175 rural men, and 150 rural women. The word rural is used to identify our non-college Ss, who were mostly from very small towns or the country and who were skewed to the right on a scale of years of formal education. Concerning the Ss as a whole, they may be described as ordinary undistinguished people living in a predominantly rural region wherein agriculture is the modal occupation and where the ratio of Negro to White is roughly one.

### C. THE INTERVIEWS

The interviews were conducted by college educated adults who were trained and supervised by a psychologist, qualified for his task by experience gained as a wage-earning interviewer in government research. Each interviewer saw only members of his own race and sex. To discover critical personal experiences involving Negro and white individuals, each interviewer followed a fixed procedure in selecting Ss and in questioning them in private. By way of establishing his identity as an authorized interviewer representing an institution of higher learning, he said,

I am a member of a group from——college (or University) that is trying to get some facts about personality and race problems. You are one of hundreds of people that I hope will help us by telling me just a few things. Your name will not be used in connection with what you say. In fact, no one will know that you tell me anything unless you tell it.

Then, following establishment of satisfactory rapport, the questions were asked as follows:

1. Some white people do things that we don't like. When you were made the maddest in your life by a white person, what had he done? What happened?



2. Now, can you tell me *another* time when you were made very mad by a white person? What had he done? What happened?

3. Which time were you *really* the maddest?

4. Some colored people do things that we don't like. When you were made the maddest in your life by a colored person, what had he done? What happened?

5. Next, can you tell me another time when you were made very mad by a colored person? What had he done? What happened?

6. Which time were you *really* the maddest?

7. When you were pleased the most in your life by a white person, what happened?

8. Can you tell me another time when you were pleased by a white person?

9. Now, can you tell me when you were most pleased by a *colored* person? (The phrase "outside your family" was used in connection with any question if the answer made it necessary.)

10. Last, can you tell me another time when a colored person pleased you very much?

11. Now, I don't need your name, but I would like to get some background information.

12. What do you really think is the most objectional thing about (*the other race*)?

Recording of replies was done during the interview and hence preceded the collection of background data regarding education, economic status, rural or urban residence, and age. Half the Ss were asked questions about their encounters with members of their own race first and then were asked about the other race. The other half of the Ss were asked first about encounters with the opposite race. This order was designed to balance possible effects of answers to first questions upon answers to subsequent questions. Whether it did so, or not, is unknown.

This paper presents results pertaining to unpleasant incidents only. The remainder of the data will appear elsewhere.

#### D. RESULTS

The most difficult task connected with the study was that of refining the thousands of answers so that they could be comprehended. The principal investigator and his chief assistant assigned categories, or class names, to all incidents that were reported. It was then possible to construct frequency tables which afforded comparisons of groups, sexes, and so on.

The manner in which classification was achieved was as follows. The first interviewee's reply to the first question was read by the two investigators. A category name was recorded and a tally checked by it. (Verbs were almost

always needed to describe the incidents because the manner of questioning had been carefully planned to yield descriptions of actions, for example, "pelting," "swindling," "going my bail.") Then the answer to the second question of the same form was read and, if appropriate, a tally was checked. If the reply was different, a new category was added and checked. Next were read the replies to the same questions by the second respondent and tallies or new categories were recorded, and so on for all the answers to all the questions by all Ss. Sixty-three categories were required for classification of the unpleasant incidents. Classification of answers of various subgroups, e.g., college Negro girls, required only a portion of the entire list of categories. Consequently, some categories have a frequency of zero for some subgroups.

Replies to the open question on objectionable things about the opposite race were classified in the same manner as were the incidents-data. There was one important difference in the outcome; few verbs were applicable as names for the objections.

The main data are summarized in Table 1, which shows the reported actions by Negroes and by whites that anger Negroes and whites. It should be noted that the data are in per cents of incidents. They are presented in this manner because two incidents were sought from each S. Actions in this table are ranked according to descending frequencies as reported by the Negroes for incidents involving both Negroes and whites. A second ranking of much of the same data, but in terms of frequencies reported by the whites, appears in Table 2. A collation of this ranking with the top ten categories in Table 1 shows for the two races six common action-categories: No incident, Criticizing severely (chewing out), Cursing or attacking a (*Negro or white*), Reneging on debts or promises, Commanding or acting arbitrarily, and Pilfering. Of these six the most frequent was No incident; many people were simply unable to recall a time when they had been angered by a member of either race. Whites exceeded Negroes in this respect. Furthermore, whites recalled fewer incidents involving Negroes than they did involving whites.

Criticizing severely, another overlapping category, was second in frequency of occurrence but it was first in producing anger since the no-incident category, although being data, was not a cause of anger. Negro reports of severe criticism outnumbered those for the whites two to one and were instigated about evenly by the two races. The whites rarely reported being criticized by the Negroes. More information about the remainder of the first ten categories may be found in later tables.

Subgroup comparisons may be made from tables constructed of various combinations of the top ten categories reported by each subgroup. (The



TABLE 1  
 REPORTED ACTIONS INVOLVING NEGROES AND WHITES THAT ANGER  
 NEGROES AND WHITES  
 (Data are in per cents of actions reported)

Action categories	Actions that anger Negroes			Actions that anger Whites		
	By Negroes	By Whites	Total	By Negroes	By Whites	Total
1. Criticizing severely (chewing out)	19.07	20.56	19.82	2.21	16.92*	9.57*
2. No incident	13.89	12.59	13.24	23.78	11.97*	17.87*
3. Cursing or attacking a Negro	8.80	7.31	8.06	.99	.76	.88*
4. Nicknaming contemptuously	.19	12.96*	6.57	—	—	—*
5. Reneging on debts or promises	6.02	2.69*	4.35	7.32	8.77	8.04*
6. Accosting a Negro Woman	2.50	4.07	3.29	—	—	—*
7. Being a "White Man's Nigger"	6.20	—*	3.10	—	—	—*
8. Commanding or acting arbitrarily	4.91	1.11*	3.01	—	8.69*	4.34
9. Being rowdy	5.09	.83*	2.96	1.60	3.43	2.52
10. Pilfering	4.72	.83*	2.78	8.84	3.81*	6.33*
11. Teasing maliciously	2.04	2.04	2.04	.76	2.90*	1.83
12. Serving another out of turn	.19	3.61*	1.90	—	—	—*
13. Advocating the "White Man's Nigger"	—	3.80*	1.90	—	—	—*
14. Driving carelessly	.83	2.31	1.57	6.94	3.66*	5.30*
15. Lying	2.87	.19*	1.53	.08	3.58*	1.83
16. Being uncivil	1.48	1.57	1.53	.94	2.44*	1.68
17. Making threats	.56	2.41*	1.48	.15	.84	.50*
18. Refusing service	—	2.87*	1.44	—	—	—*
19. Making an immoderate request	.83	1.85	1.34	.38	.99	.69
20. Working carelessly	1.39	.93	1.16	10.67	3.35*	7.01*
21. Being a "Stool Pigeon" for Whites	2.22	—*	1.11	—	—	—*
22. Giving less than the promised pay	.09	2.04*	1.06	—	.76*	.38
23. Pelting	—	1.94*	.97	—	—	—*
24. Cursing before Negro Women	1.20	.46	.83	—	—	—*
25. Being snobbish	1.30	.37	.83	—	1.45*	.72
26. Mistrusting (watching closely)	.28	1.39	.83	.15	.84	.50
27. Swindling	.56	1.02	.79	.08	.69	.38
28. Urging doggedly	—	1.57*	.79	.30	1.37*	.84
29. Giving unsolicited advice	.65	.83	.74	—	—	—*
30. Gossiping	1.39	—*	.69	.08	.69	.38

TABLE 1 (continued)

Action categories	Actions that anger Negroes			Actions that anger Whites		
	By Negroes	By Whites	Total	By Negroes	By Whites	Total
31. Destroying property	1.02	.28	.65	.30	1.83*	1.07
32. Meddling	.83	.46	.65	1.07	.61	.84
33. Grumbling	—	1.20*	.60	5.79	—*	2.90*
34. Using property without permission	1.11	.09*	.60	1.30	1.52	1.41
35. Neglecting his family	1.11	.09*	.60	.46	.15	.30
36. "Hogging" the sidewalk	—	1.11*	.56	1.68	—*	.84
37. Mistreating pets	—	1.11*	.56	.15	—	.08
38. Interfering with dating	1.02	—*	.51	—	3.05*	1.52*
39. Attempting to steal job	.74	.19	.46	—	.84*	.42
40. Pandering	.93	—*	.46	—	—	—*
41. Giving low grades	.93	—*	.46	—	—	—*
42. Dressing gaudily	.65	—	.32	—	—	—
43. Getting ahead of turn	—	.56	.28	1.22	.76	.99*
44. Accosting a white man	.46	—	.23	—	.23	.11
45. Releasing stock	.46	—	.23	—	.46	.23
46. Requesting pandering	—	.46	.23	—	—	—
47. Cursing or attacking a white	.37	—	.19	5.11	9.22*	7.16*
48. Accosting a Negro man	.37	—	.19	—	.08	.04
49. Rummaging	.37	—	.19	.15	—	.08
50. Giving off an offensive odor	.19	—	.09	.23	.08	.15
51. Releasing pets or children	—	.19	.09	.08	.46	.27
52. Prowling	—	.09	.05	.99	.15	.57*
53. Attempting to integrate	.09	—	.05	.08	—	.04
54. Begging	.09	—	.05	.84	.23	.53*
55. Cursing before white women	—	—	—	1.07	1.14	1.11*
56. Addressing by first name	—	—	—	1.52	.08*	.80*
57. Accosting a white woman	—	—	—	4.19	.61*	2.40*
58. Acting as a social equal	—	—	—	5.18	—*	2.59*
59. Using "White only" facilities	—	—	—	.99	—*	.50*
60. Imitating	—	—	—	—	.15	.08
61. Mistreating a friend	—	—	—	—	.38	.19
62. Stealing maids	—	—	—	—	.08	.04
63. Being clannish	—	—	—	2.36	—*	1.18*
Total per cent	100.010	99.980	100.010	100.030	100.020	100.030
Total incidents	1080	1080	2160	1312	1312	2624

\* Critical ratio of three or more for the difference between the number marked and its counterpart.



TABLE 2

THE TOP TEN REPORTED ACTIONS INVOLVING NEGROES AND WHITES THAT ANGER WHITES  
(Data from Table 1)

Actions reported	Actions that anger Whites		
	By Negroes %	By Whites %	Total %
1. No incident	23.78	11.97*	17.87
2. Criticizing severely (chewing out)	2.21	16.92*	9.57
3. Reneging on debts or promises	7.32	8.77	8.04
4. Cursing or attacking a white	5.11	9.22*	7.16
5. Working carelessly	10.67	3.35*	7.01
6. Pilfering	8.84	3.81*	6.33
7. Driving carelessly	6.94	3.66*	5.30
8. Commanding or acting arbitrarily	—	8.69*	4.34
9. Grumbling	5.79	—*	2.90
10. Acting as a social equal	5.18	—*	2.59
Total per cent	75.84	66.39	71.11
Total incidents = 2624			

\* Critical ratio of three or more for the difference between the number marked and its counterpart.

remaining subgroup categories, all of which have small frequencies, appear as group data in Table 1.)

Table 3, containing data from the four subgroups of men, was constructed in two parts. In Part I, the first column on the left is a list of nineteen actions of Negroes; the number of action-categories necessary to include the top ten categories of each of the four subgroups. The second column shows frequencies for the nineteen categories in per cents of total Negro-actions reported by the rural Negro subgroup. The meaning of a number in this column may be seen by noting, say, the number 15.00, the second entry, which indicates that 15 per cent of the Negroes' actions that angered rural Negro men were "renegings on debts," etc. A dash (—) in the column indicates a frequency of zero for its corresponding category. The third column was constructed in exactly the same way as the second column except that the data were taken from the college Negro subgroups. The next column, total, is for the rural and college Negroes combined. The first entry in the column, 22.20, shows the per cent of male Negro reports that fell in the category of no incidents.

The data for Negro-actions that angered the rural and college white subgroups are tabulated in the next three columns. The last column, headed "All Men," is the sum of the "Total" columns for Negroes and whites. It shows the per cent of men's responses, Negroes and whites, that fell into each Negro-action category. Part II of Table 3 shows actions by whites that

TABLE 3  
ACTIONS BY NEGROES AND WHITES THAT ANGER NEGRO MEN AND WHITE MEN  
(Data are per cents of all actions reported by men)

I. Negroes' actions that anger men	Rural	Negro College	Total	Rural	White College	Total	All Men
No incident	36.67	6.25	22.20	14.86	22.22	18.59	20.20
Reneging on debts or promises	15.00	2.94	9.27	11.43	4.17	7.75	8.42
Criticizing severely (chewing out)	3.67	23.53	13.11	6.00	1.39	3.66	7.88
Working carelessly	.33	1.84	1.05	10.29	10.00	10.14	6.08
Pilfering	3.33	5.51	4.37	7.71	5.56	6.62	5.62
Driving carelessly	2.00	.37	1.22	9.71	7.22	8.45	5.23
Cursing or attacking white	.33	.74	.53	7.71	8.33	8.03	4.68
Cursing or attacking a Negro	6.33	8.82	7.52	2.57	.56	1.55	4.21
Being rowdy	4.00	10.66	7.71	1.71	1.94	1.83	4.21
Acting as a social equal	—	—	—	4.29	10.28	7.32	4.06
Being a "White Man's Nigger"	7.33	10.66	8.92	—	—	—	3.98
Grumbling	—	—	—	3.43	9.17	6.34	3.51
Being clannish	—	—	—	2.86	3.33	3.10	1.72
Commanding or acting arbitrarily	1.33	6.62	3.85	—	—	—	1.71
Accosting a white woman	—	—	—	2.57	3.61	3.10	1.71
Being a "stool pigeon" for whites	2.33	4.41	3.32	—	—	—	1.48
Addressing by first name	—	—	—	2.86	—	1.41	.78
Pandering	1.00	2.57	1.75	—	—	—	.78
Neglecting family	2.00	—	1.05	.29	.28	.28	.63
Total per cent	85.65	84.92	85.33	88.29	88.06	88.17	86.89
Total incidents	300	272	572	350	360	710	1282

angered men. The data are presented in the same way as are those in Part I of the table.

Table 4, which presents the data for the four female subgroups, was constructed in the same manner as was Table 3. In Part I, the first column is a list of 24 Negro-action categories; the number necessary to include the top ten categories for the four subgroups. The second column shows the frequencies, in per cents, with which each of the twenty-four categories in column one was reported by the rural Negro women. The second entry in the column indicates that 31.82 per cent of all rural Negroes' incidents



TABLE 3 (continued)

II. Whites' actions that anger men	Rural	Negro College	Total	Rural	White College	Total	All Men
Criticizing severely (chewing out)	16.67	24.26	20.23	16.29	16.94	16.62	18.25
No incident	30.00	4.78	18.01	9.71	8.06	8.07	12.95
Cursing or attacking a white	—	—	—	14.00	14.17	14.08	7.80
Nicknaming contemp- tuously	9.67	18.75	13.99	—	—	—	6.24
Commanding or acting arbitrarily	.67	1.84	1.22	8.57	11.39	10.00	6.09
Reneging on bets or promises	3.33	1.84	2.62	12.57	5.28	8.87	6.08
Cursing or attacking a Negro	6.00	13.97	9.79	.29	.28	.28	5.51
Teasing maliciously	3.67	1.47	2.62	3.71	4.44	4.08	3.43
Driving carelessly	3.33	3.31	3.32	3.14	3.05	3.10	3.20
Being rowdy	.33	2.21	1.22	4.86	3.61	4.23	2.89
Pilfering	.33	.37	.35	5.14	3.33	4.23	2.50
Interfering with dating	—	—	—	1.14	7.50	4.36	2.42
Making threats	2.67	3.68	3.15	1.43	.56	.99	1.95
Working carelessly	—	.74	.35	4.29	.22	3.24	1.95
Lying	—	.74	.35	2.00	4.17	3.10	1.87
Accosting a Negro woman	1.33	3.31	2.27	—	—	—	1.71
Urging doggedly	.33	2.21	1.22	1.43	2.22	1.83	1.56
Advocating the "White Man's Nigger"	2.00	4.41	3.15	—	—	—	1.41
Refusing service	1.33	4.41	2.80	—	—	—	1.25
Serving another out of turn	3.33	1.10	2.27	—	—	—	1.01
Mistrusting (watching closely)	2.33	.37	1.40	—	—	—	.63
Total per cent	87.32	93.77	90.38	88.57	87.22	87.08	90.72
Total incidents	300	272	572	350	360	710	1282

involved criticism by Negroes. The third column of the table is a list of the frequencies of Negro-actions that angered the college Negro group. The fourth column, "Total," is a combination of frequencies of the rural and college Negro subgroups. The next three columns show for the two white subgroups the frequencies of Negro-instigated reports. The final column gives, for all women combined, the ranks merited by the twenty-four categories.

TABLE 4  
ACTIONS BY NEGROES AND WHITES THAT ANGER NEGRO WOMEN AND WHITE WOMEN  
(Data are per cents of all actions reported by women)

I. Negroes' actions that anger women	Rural	Negro College	Total	Rural	White College	Total	All Wom- en
No incident	1.82	9.55	4.53	16.33	43.38	29.90	18.29
Criticizing severely (chewing out)	31.82	14.61	25.79	—	.99	.50	12.07
Pilfering	5.45	4.49	5.12	17.33	5.63	11.46	8.56
Working carelessly	1.52	2.25	1.77	16.33	6.29	11.30	6.94
Cursing or attacking a Negro	10.30	10.11	10.24	—	.66	.33	4.86
Reneging on debts or promises	3.33	.56	2.36	12.67	.99	6.81	4.77
Driving carelessly	—	1.12	.39	5.67	4.64	5.15	2.97
Accosting a white woman	—	—	—	1.33	9.60	5.48	2.97
Commanding or acting arbitrarily	5.15	7.87	6.10	—	—	—	2.79
Grumbling	—	—	—	5.67	4.64	5.15	2.79
Lying	4.24	6.74	5.12	—	.33	.17	2.43
Accosting a Negro woman	3.64	5.62	4.33	—	—	—	1.98
Being rowdy	2.42	3.37	2.76	2.00	.66	1.33	1.98
Teasing maliciously	4.55	1.12	3.35	1.00	.33	.66	1.89
Being uncivil	1.21	4.49	2.36	1.00	1.66	1.33	1.80
Meddling	1.21	.56	.98	3.00	1.32	2.16	1.62
Being a "White Man's Nigger"	3.03	3.37	3.15	—	—	—	1.44
Acting as a social equal	—	—	—	2.33	2.98	2.66	1.44
Being snobbish	3.64	1.12	2.76	—	—	—	1.26
Addressing by first name	—	—	—	1.00	2.32	1.66	.90
"Hogging the side- walk"	—	—	—	1.33	1.99	1.66	.90
Being clannish	—	—	—	—	2.98	1.49	.81
Prowling	—	—	—	2.00	.99	1.49	.81
Giving low grades	—	3.93	1.38	—	—	—	.63
Total per cent	83.33	80.88	82.49	88.99	93.38	90.69	86.90
Total incidents	330	178	508	300	302	602	1110

Part II of Table 4 shows a listing of twenty-six actions by whites that include the top ten most frequently reported action-categories for the four female subgroups. Arrangement of the columns duplicates that of Part I.

The classes of responses to the question, "What do you really think is the



TABLE 4 (continued)

II. Whites' actions that anger women	Rural	Negro College	Total	Rural	White College	Total	All Wom- en
Criticizing severely (chewing out)	25.15	12.92	20.87	20.00	14.57	17.28	18.92
No incident	2.42	14.05	6.50	3.33	27.81	15.61	11.44
Reneging on debts or promises	3.33	1.69	276	9.00	8.28	8.64	5.95
Nicknaming contemp- tuously	10.61	14.05	11.81	—	—	—	5.41
Commanding or acting arbitrarily	1.52	—	.98	5.33	8.94	7.14	4.32
Advocating the "White Man's Nigger"	3.64	6.18	4.53	—	—	—	3.15
Being uncivil	2.12	2.81	2.36	5.00	1.99	3.49	2.97
Driving carelessly	.91	1.69	1.18	7.00	1.66	4.32	2.88
Cursing or attacking a Negro	2.73	7.87	4.53	—	2.65	1.33	2.79
Accosting a Negro woman	5.76	6.74	6.10	—	—	—	2.79
Working carelessly	1.82	1.12	1.58	5.33	1.66	3.49	2.61
Pilfering	1.52	1.12	1.38	5.33	1.32	3.32	2.43
Serving another out of turn	3.03	8.99	5.12	—	—	—	2.34
Lying	—	—	—	4.33	3.97	4.15	2.25
Making an immod- erate request	2.42	2.25	2.36	1.67	1.99	1.83	2.07
Meddling	3.64	1.12	2.76	2.33	.33	1.33	1.98
Cursing or attacking a white	—	—	—	2.67	4.30	3.49	1.89
Mistrusting (watching closely)	.91	2.25	1.38	1.33	2.32	1.83	1.62
Pelting	4.55	1.69	3.54	—	—	—	1.62
Giving less than the promised pay	3.94	1.69	3.15	.67	—	.33	1.62
Teasing maliciously	1.82	.56	1.38	.67	2.32	1.49	1.44
Refusing service	2.12	4.49	2.95	—	—	—	1.35
Being snobbish	.61	—	.39	1.00	3.31	2.16	1.35
Destroying property	.61	—	.39	3.33	—	1.66	1.08
"Hogging the side- walk"	2.12	2.25	2.17	—	—	—	.99
Interfering with dating	—	—	—	.67	2.32	1.49	.81
Total per cent	87.30	95.53	90.17	78.99	89.74	84.38	88.07
Total incidents	330	178	508	300	302	602	1110

TABLE 5  
INTERRACIAL OBJECTIONS

I. Negroes' objections to whites	Rural	Women College	Total	Rural	Men College	Total	All Negroes
Presumption of superiority	29.10	55.06	38.19	47.33	38.97	43.35	40.93
Oppression of Negroes	20.60	19.10	20.08	41.33	23.53	32.86	26.85
Prejudice	8.48	17.98	11.81	4.00	19.85	11.54	11.67
No data	7.88	3.37	6.30	1.33	9.56	5.24	5.74
One-way sex relations	7.27	—	4.72	3.33	4.41	3.85	4.26
Insincerity	11.51	—	7.48	—	—	—	3.52
Saying "Nigger"	3.03	2.25	2.76	.67	2.21	1.40	2.04
One-way "cussing"	4.24	—	2.76	2.00	—	1.05	1.85
Dislike of Negroes' color	5.45	—	3.54	—	—	—	1.67
"No objection"	2.42	—	1.57	—	1.47	.70	1.11
Color	—	1.12	.39	—	—	—	.19
"Nasty" personalities	—	1.12	.39	—	—	—	.19
II. Whites' objections to Negroes	Rural	Women College	Total	Rural	Men College	Total	All Whites
Immorality	67.99	5.30	36.54	8.00	13.89	10.99	22.71
Uncleanliness	12.00	12.58	12.29	17.71	8.89	13.24	12.80
Trying to integrate	1.33	5.96	3.65	22.86	13.33	18.03	11.43
Indecency	—	19.20	9.63	7.43	11.67	9.58	9.60
Ignorance	1.33	11.26	6.31	8.00	14.44	11.27	8.99
Indolence	2.67	3.97	3.32	7.43	12.78	10.14	7.01
Presumption of equality	1.33	4.64	2.99	6.86	7.22	7.04	5.18
"No objection"	.67	9.93	5.32	2.26	2.78	2.54	3.81
Irresponsibility	.67	2.65	1.66	6.86	3.89	5.35	3.66
Arrogance	—	5.96	2.99	4.57	3.33	3.94	3.51
Odor	.67	4.64	2.66	1.71	5.55	3.66	3.20
No data	2.67	4.64	3.65	.57	1.67	1.13	2.29
Aggression	—	5.96	2.99	2.29	—	1.13	1.98
Lying	4.00	—	1.99	2.29	—	1.13	1.52
Stealing	3.33	—	1.66	.57	—	.28	.91
Color	—	3.31	1.66	—	.55	.28	.91
Subservience	.67	—	.33	—	—	—	.15
Drinking	.67	—	.33	—	—	—	.15
Clannishness	—	—	—	.57	—	.28	.15

most objectionable thing about the (*other*) race?" appear in Table 5. Part I of the table shows in the first column a list of the Negroes' objections to whites. The second column shows the per cents of rural Negro females' re-



sponses per objection. The other columns in the tables are self-explanatory. Part II of the table shows the responses of the whites to the same question.

Table 6 shows the relationship found between objections to the opposite race and incidents of friction with individuals of that race. The top row

TABLE 6  
RELATIONSHIPS BETWEEN INCIDENTS OF FRICTION AND RACIAL OBJECTIONS

	Negroes				Whites			
	Men		Women		Men		Women	
	Rural	College	Rural	College	Rural	College	Rural	College
Number interviewed	150	136	165	89	175	180	150	151
Relationships	23	42	30	23	4	18	9	0
Per cent	15.33	30.88	18.18	25.84	2.29	10.00	6.00	—

indicates the number of people in the subgroups. The middle row indicates the total number of times that those people expressed an objection that was in harmony with an incident of friction that they had reported. The last row shows the same numbers in per cents; thus, 15.33 per cent of the rural Negro men expressed objections that might have been influenced by frictional contacts with white individuals.

#### E. SOME COMMENTS

Our respondents had a strong inclination to describe their encounters with the other race not as personal experiences but as stereotyped generalizations. The words "racial problems" mentioned by our interviewers in their introductory statements acted at once as stimuli for such comments as "God made 'em that way," "Ham," "Our girls are like pigs—scratch their sides and they flop over on their backs," "No Negro civilization since the Pharaohs," and the like. Such words or thoughts by Ss may have blocked their attempts at recall of specific incidents. On the other hand, difficulty in recall of specific incidents may have been due to the absence of frictional contacts; that is, the two races may live pretty much as do the owls and snakes and prairie dogs, and, for that matter, all other animals that respect through learning of a positive sort the poetically called "territorial rights" of others. Of course, the gimmick in the owl-snake-dog story is the live-happily-together part; the happiness is not measured. In our study the gimmick may be the hate-the-other-race assumption. If genuine hate could be produced only by personal incidents that involve intense friction there just shouldn't be very much hate. There weren't enough incidents to go round. Furthermore, the highest cate-

gory of reported encounters was Criticising severely, which in itself is not really damaging, not pain producing.

It might be argued that our Ss were too shrewd to reveal their own involvement in the "touchy" racial issue. To subscribe to this notion, however, would require overlooking too many Ss who recalled no personal encounters of aversive emotional nature with the other race yet who had no reservations about expressing adverse opinions of it.

We recognize and regret that in categorizing the raw data we caused a loss of considerable information. For example, here are two of the kind of attention-holding incidents (thematic stimuli?) whose pathos does not appear in the action categories. One involved a Negro woman who on leaving a store had the door politely held open by an elderly white man until his eyes encountered her color whereupon he instantly released the door which struck the woman head-on while he meanwhile averted his gaze and vanished in the crowded street. The other example concerned a young white plantation laborer who left his tractor seat to spend a few minutes rest with a Negro couple that was chopping behind him. When he commented about the hoeing, the woman, thinking that he was speaking about her, called him a "red-necked bastard."

#### F. CONCLUSIONS

This study was carried out so that a list of actions by each race that annoy the other race could be constructed. Such lists are in Tables 1 and 2. Is there a good purpose to which they could be put? If whites had refrained from indulging in the apparent trivialities of chewing out, cursing, and nick-naming contemptuously ("boy," "auntie," etc.) our Negro Ss, over half of their incidents-reports would have been different. And if Negroes had not pilfered, worked carelessly, and reneged on petty debts and promises, more than half the incidents-data from our white subjects would have been altered. Perhaps the lists could be educationally useful.

#### REFERENCES

1. DAVIS, A., GARDNER, B. B., & GARDNER, M. *Deep South: A Social Anthropological Study of Caste and Class*. Chicago, Ill.: Univ. Chicago Press, 1941.
2. POWDERMAKER, H. *After Freedom: A Cultural Study in the Deep South*. New York: Viking, 1939.

*Department of Psychology*  
*University of Mississippi*  
*University, Mississippi*



## STEREOSCOPIC DEPTH PHENOMENA WITH WITTE-KÖNIG EFFECT\*

Ottawa, Canada

---

J. L. ZAJAC<sup>1</sup>

---

The stereoscopic fusion of segments of lines of unequal length yields, as is well known, a depth effect. For example, if a shorter line is placed in front of the left eye and a longer one in front of the right eye, and a fixation-point is chosen between *O* and the lines (with uncrossed, natural stereoscopy), an image results whose left end is seen farther away from *O* and right end nearer to *O*. With crossed fusion, i.e., when the fixation-point is situated behind the lines, the relief of the stereoscopic image is reversed and the fused image is larger. The stereoscopic image as well as monocular images (not represented in Figure 1) seen on each side of the stereoscopic image, look smaller than the corresponding segments in the stimulus-figure itself, but with crossed fusion they look larger. All these phenomena are in conformity with the laws of binocular correspondence and disparity and can be adequately illustrated by the scheme of intersection of directional visual lines shown in Figure 1. When the segments are of equal length, no such depth effect occurs, but with the uncrossed fusion the stereoscopic image is still shorter, and with crossed fusion larger, than the actual lines.

When two short lines separated by a gap are placed in front of one eye and to the other eye is presented only one segment not very different in length from the total length of the other two, some phenomena might be observed, named by Helson and Wilkinson the Witte-König paradoxical fusion effect (1). The shorter the gap between the two segments, the more frequent and more steady is the maintenance of the gap in the stereoscopic image, and when some lower limit for the length of the gap is reached, the closure never occurs.<sup>2</sup> Helson and Wilkinson characterize this phenomenon as due to contrast.

The author has observed and analyzed phenomena of depth resulting

---

\* Received in the Editorial Office on April 11, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author is indebted to Dr. Catherine J. D. Jarvis of the National Research Council of Canada for revising and correcting the English of this paper.

<sup>2</sup> The author made the corresponding observations independently of Witte-König's findings, when carrying experiments with some stereoscopically fused drawings by Linschoten, Werner, Koffka and his own.

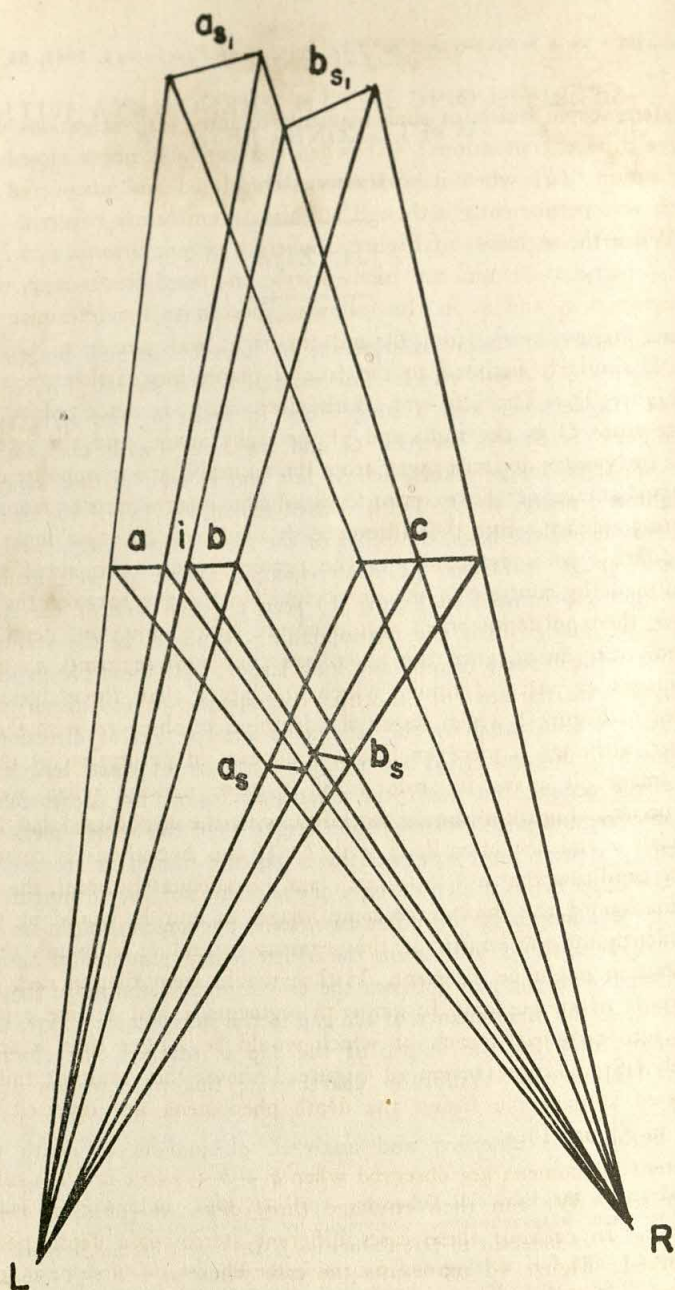


FIGURE 1

Crossed and uncrossed stereoscopic fusion of  $a$  with half of  $c$ , and of  $b$  with other half of  $c$ . L—left eye, R—right eye,  $a$ ,  $b$ ,  $c$ —segments of lines,  $i$ —gap between  $a$  and  $b$ ,  $a_s$ ,  $b_s$ —stereoscopic images from uncrossed fusion,  $a_{s1}$ ,  $b_{s1}$ —from crossed fusion.



from stereoscopic fusion of such segments of lines. Investigations were made in three different situations: (a) when the gap was never closed in stereoscopic fusion; (b) when it was alternately closed and uncovered; and (c) when it was permanently although not always uniformly covered.

1. When the segments of Figure 2 where  $a = b = 9$  mm.,  $i = 2$  mm. and  $a + b + i = c = 20$  mm. are fused in free uncrossed stereoscopy, we observe two segments  $a_s$  and  $b_s$  in the following positions: the left image with its left end farther away from  $O$  and the right end nearer to  $O$ , the right image  $b_s$  similarly inclined to the frontal plane, but farther away from  $O$  (see Figure 1). The left end of the left image is perceived at the same distance from  $O$  as the right end of the right image, and the interior ends of both images are declined away from the frontal plane in opposite directions. The equidistance of the exterior ends of the images results from the law of correspondence. But the difference in depth of the two inner edges is more difficult to understand. If the perceived gap represented the interruption made by contrast in the line  $c$  equal to the gap between the segments  $a$  and  $b$ , then no depth effect would occur. Since, however, depth effect is observed, one can advance the hypothesis that both segments  $a$  and  $b$  fuse with images of parts of line  $c$  which are larger than themselves. This is indicated in Figure 1 which shows the depth effect observed with the hypothesis, that, although a piece of line  $c$  disappears in perception of the stereoscopic image, it plays its role as if it were present when stereoscopic fusion occurs. Figure 1 shows the effect with the hypothesis that half of  $c$  fuses with  $a$  and the other half with  $b$ . If this hypothesis is correct, then with the condition that  $a + b + i = c$  and  $i$  is adequately small, the direction of inclination of the two stereoscopic images  $a_s$  and  $b_s$  would be the same independently of the length of the segments  $a$  and  $b$ , although the angles of inclination might be different. It then would seem that in each case, for the purpose of stereoscopic fusion with segments  $a$  and  $b$ , line  $c$  would be divided into two parts, each of which would be longer than  $a$  and  $b$  respectively. Stereoscopic fusion of Figure 3 shows the result of this. With the crossed stereoscopic fusion the depth phenomena are reversed (see  $a_{s1}$  and  $b_{s1}$  in Figure 1).

Different phenomena are observed when  $a + b + i \neq c$  and  $i$  is sufficiently small in size. We can then envisage three cases:  $a + b < c$ ,  $a + b = c$ ,  $a + b > c$ . In each of these cases different stereoscopic depth phenomena are observed. Figure 4d represents the case where  $a + b = c$ ; in this case with uncrossed fusion two horizontal segments are seen, but the left segment is nearer to  $O$  and the right one farther away from  $O$ . Figure 4e illustrates

the case where  $a + b < c$ ; in this case, the stereoscopic image  $a_s$  is seen with its left end farther away from  $O$  and the image  $b_s$  is slightly inclined in the same direction. Figure 4f represents the case where  $a + b > c$ ; in this case the longer segment  $a_s$  is inclined to the frontal plane in the direction opposite to that in the previous case and  $b_s$  is inclined in the same direction as  $a_s$  and farther away from  $O$ .

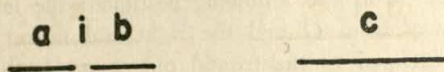


FIGURE 2

SEGMENTS OF LINES DESTINED FOR STEREOSCOPIC FUSION; DENOTATIONS AS IN FIGURE 1



FIGURE 3

Segments showing by fusion different reliefs of stereoscopic images when the lengths of  $a$  and  $b$  vary, but the size of  $i$  and  $c$  remains unchanged and  $a + b + i = c$ .

2. When the gap  $i$  between  $a$  and  $b$  surpasses certain limits, then with stereoscopic fusion the gap is intermittently closed and uncovered. We can distinguish here several phases: (a) The gap is uncovered; in this phase phenomena similar to those described above are observed. The depth effect might be enhanced in comparison with corresponding cases in Section 1, which can be understood from the fact that there are greater differences in length between  $a$  and  $b$  and corresponding fusing parts of  $c$ . (b) The gap between  $a_s$  and  $b_s$  is completely and uniformly closed: the line then appears as a straight whole line, horizontal when  $a + b + i = c$ , inclined with the left edge farther away from  $O$  when  $a + b + i < c$ , and with the left edge nearer to  $O$  if  $a + b + i > c$ . (See and fuse Figure 5def.). With crossed fusion the inclination in Figures 5e and 5f are reversed. (c) The complete and uniform closure of the gap lasts for some time and then the closing parts of line  $c$  start to fade away; this fading starts at the ends of images  $a_s$  and  $b_s$  adjoining the gap.<sup>3</sup> The central part of the closure is still maintained for some time after the parts adjoining the interiors of  $a_s$  and  $b_s$  have already disappeared. This illustrates the fact that the contrast effect is strongest at these points and can be considered as a partial explanation of

<sup>3</sup> If the lines are drawn on a white paper, the parts adjoining the interior ends of  $a_s$  and  $b_s$  do not show the brightness of the paper itself, but the brightness of an after-image, i.e., they are more brilliant than the paper; sometimes brilliant lines fill up the whole gap, being after-images of suppressed parts of line  $c$ .



the complete absence of the closure when the gap is small. When the closure starts to fade away, the uniform-line character of the stereoscopic image also disappears; the line becomes broken and appears in the relief described in 2a. Therefore, the narrower the gap the less frequently does the closure occur.<sup>4</sup>

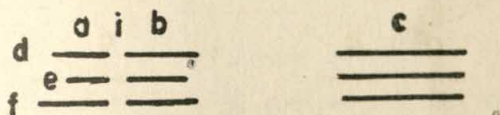


FIGURE 4

Different depth effects when  $i$  and  $c$  remain unchanged while  $a$  and  $b$  vary in length, and  $a + b = c$  (case d) or  $a + b \neq c$  (cases e and f).

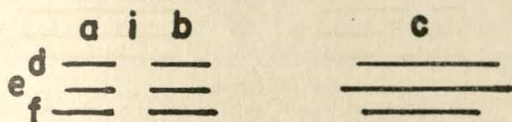


FIGURE 5

Different depth effects, when gap  $i = 5$  mm., and the relation of the length of  $a + b$  to  $c$  varies.

3. When the gap is still greater, then the closure of the gap  $i$  in the stereoscopic image is steady, although not always uniform and of a fully satiated form. The closure is rather steady but faded where the contrast effect is strongest, i.e., at the interior edges of  $a_s$  and  $b_s$ . Thus, depth phenomena are observed; the line appears to be broken into parts showing various inclinations to the frontal plane as in phase 2c above.

These phenomena in the author's opinion present a new, and so far not sufficiently investigated, aspect of stereoscopic depth perception. A thorough investigation of the whole subject would seem to be both desirable and promising.

#### REFERENCE

1. HELSON, H., & WILKINSON, A. E. A study of the Witte-König paradoxical fusion effect. *Amer. J. Psychol.*, 1958, **71**, 316-320.

1917 Saunderson Drive  
Ottawa 1, Canada

<sup>4</sup> This narrowing of the gap  $i$  can be brought about by drawing the gap smaller or by increasing the distance of the figure from  $O$ .

## TIME JUDGMENT AND SCHIZOPHRENIA: PSYCHOPHYSICAL METHOD AS A RELEVANT CONTEXTUAL FACTOR\*<sup>1</sup>

*Baylor University College of Medicine, Houston State Psychiatric Institute,  
and Vanderbilt University*

---

FRANCES R. WEBSTER, SANFORD GOLDSTONE, AND WARREN W. WEBB

---

### A. PURPOSE

Although several studies (2, 3, 4) demonstrated that schizophrenic Ss overestimated the duration of short auditory inputs, another investigation (4) showed that this overestimation was not independent of the experimental context. When ascending and descending series of arithmetically spaced durations were judged as *more* or *less* than one second, patients identified shorter sounds as equivalent to their concept than did healthy controls (2, 3, 4). This method with a geometric stimulus series did not show greater overestimation by schizophrenic Ss (4). The interpretation of the patient-control difference (2, 3) as a general disturbance in time judgment was obviously premature.

This study altered another property of the experimental context, the psychophysical method, hypothesizing that the previously used modified method of limits procedure (2, 3, 4) was a more complex judgment task than the more restricted and defined constant method. There is a greater conceptual requirement placed upon S in a situation where he selects his own range of input and where previous responses determine the following stimuli, than within a fixed and invariant context. If complexity of the task is a relevant variable in producing the patient-control difference (2, 3, 4), this difference should be attenuated with the constant method. Additionally, the hypothesis was explored that schizophrenics would use fewer judgment categories than control Ss because of the increased complexity involved in employing a magnitude continuum with many units.

### B. METHOD

#### 1. Procedure

An electronic timer activated auditory durations with a reliable and accurate range of .01 to 9.99 seconds. An audio-oscillator and headphones

---

\* Received in the Editorial Office on April 24, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Supported by U.S.P.H.S. Grant M-1121 and the Hogg Foundation.



provided a tone of 725 cps and comfortable amplitude. Ss made single stimulus judgments of a randomized series of seven durations, .15, .45, .75, 1.05, 1.35, 1.65, 1.95 seconds, along the following nine category scale: 1 *very much less than a clock second*; 2 *much less*; 3 *less*; 4 *slightly less*; 5 *equal*; 6 *slightly more*; 7 *more*; 8 *much more*; 9 *very much more than a clock second*. Ss responded with numbers and were permitted to use more extreme categories. Each duration was judged ten times.

## 2. Subjects

The following populations were studied: *Healthy*, (H).<sup>2</sup> Twenty paid college student volunteers. *Schizophrenic*, (S).<sup>3</sup> Twenty hospitalized patients whose primary diagnosis was schizophrenia. *Other Psychiatric* (P). Twenty non-psychotic psychiatric patients mostly classified as anxiety reaction or depressive reaction. *Physically Disabled* (D). Twenty patients from orthopedic, surgical and physical medicine services without a psychiatric diagnosis.

## C. RESULTS

Ss judged the series of seventy durations and the mean response to each duration was obtained. A trend test based upon analysis of variance compared group curves. *Populations* was the between and *Durations* the within Ss factors.

Figure 1 shows the curves for H, S, P, and D. Although the trend test indicates a significant effect due to *Populations* ( $F\ 3/76 = 3.11, P < .05$ ), the *Populations*  $\times$  *Durations* interaction ( $F\ 18/456 = 3.74, P < .001$ ) shows that the predominant difference is in slope.

Inspection of Figure 1 and the analysis indicate no increased overestimation by S. The patient-control difference found with another psychophysical method (2, 3, 4) was eliminated.

Slope differences indicate the range of categories used by a population. A steep slope is associated with a wider range of categories. Figure 1 shows the steepest slope for H, and the greatest apparent slope difference between H and S.

The slope differences were studied by comparing pairs of populations with these results: Significant *Population*  $\times$  *Durations* interactions were found when H was compared with S ( $F\ 6/228 = 10.07, P < .001$ ), P

<sup>2</sup> Hereafter, the alphabetic coding H, S, P, and D, will be used to designate populations.

<sup>3</sup> All patients were hospitalized in the Houston Veterans Administration Hospital. The authors are indebted to A. Pokorny, M.D., S. Cleveland, Ph.D. and the staff for their cooperation.

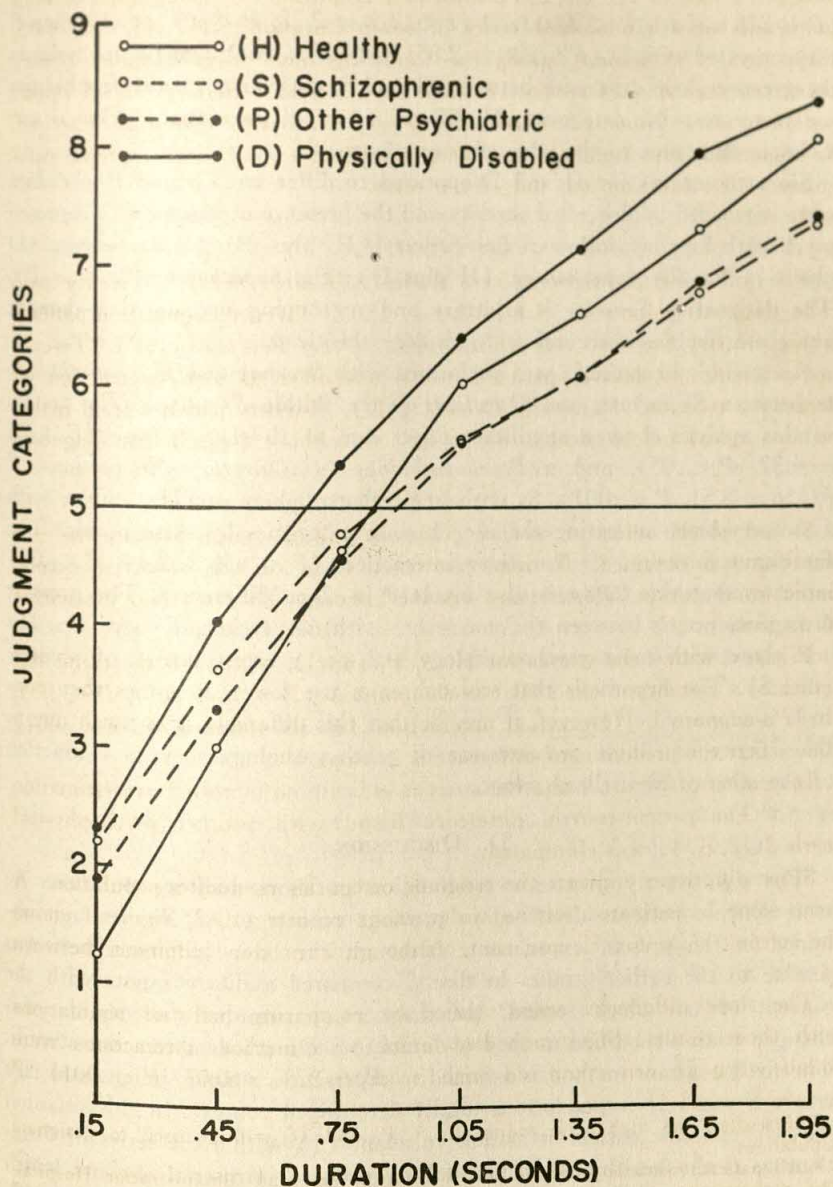


FIGURE 1  
GROUP CURVES FOR THE HEALTHY (H), SCHIZOPHRENIC (S), OTHER PSYCHIATRIC (P),  
AND PHYSICALLY DISABLED (D) POPULATIONS. EACH POINT IS THE MEAN RESPONSE  
OF THE 20 Ss IN THE GROUP TO A GIVEN DURATION.



( $F 6/228 = 6.61, P < .001$ ), D ( $F 6/228 = 2.28, P < .05$ ), and, when S was compared with D ( $F 6/228 = 2.85, P < .05$ ). Although Figure 1 shows the greatest slope difference between H and S suggesting that schizophrenics use fewer response categories, H differed from all populations. This is not a simple effect due to the schizophrenia alone.

Since the curves for H and D appeared to differ from S and P, the data were regrouped by degree of severity and the presence or absence of diagnosed psychopathology as follows: *Less-Severe*, (H plus P); *More-Severe*, (D plus S); *No Psychopathology*, (H plus D); *Psychopathology*, (S plus P). The designation *Severity* is arbitrary and regrouping was an after-thought but these liberties were taken to explore this lead.

An analysis of variance was performed with *Severity* and *Psychopathology* as between Ss factors, and *Durations* as the within Ss factor. The results of this analysis show a significant effect due to *Psychopathology* ( $F 1/76 = 4.32, P < .05$ ), and a *Psychopathology*  $\times$  *Durations* interaction ( $F 6/456 = 8.51, P < .001$ ). Ss without psychopathology provided curves with a steeper slope, indicating the use of more categories by these people. The significant *Severity*  $\times$  *Durations* interaction ( $F 6/456 = 2.44, P < .05$ ) indicates that this factor is also involved in slope differences. The greatest divergence occurs between the non-severe with no psychopathology (i.e., H) and those with most psychopathology and the greatest degree of severity (i.e., S). The hypothesis that schizophrenics use fewer categories than controls is supported. However, it may be that this difference is as much due to the severity of illness and presence of psychopathology as it is a function of the schizophrenic illness alone.

#### D. DISCUSSION

The significantly greater overestimation of short auditory durations by schizophrenic patients described in previous reports (2, 3, 4) was not obtained in the present experiment. Although the basic judgment task was similar to the earlier studies in that S compared auditory inputs with the concept of one clock second, the manner of stimulus presentation was changed from a modified method of limits to the method of constant stimuli. The fixed constant method is assumed to represent a simpler judgmental task where S passively responds to a totally determined context. In this situation the schizophrenic patient is as able to adapt (1) within the series of durations as is the healthy S without falling prey to intraserial anchors which dominate an arithmetic interstimulus interval (4). On the other hand, with the modified method of limits procedure (2, 3, 4) S does not respond

passively to a fixed stimulus series but instrumentally determines the size of immediate inputs and the range of the overall series. Under this less passive, more instrumental condition, the schizophrenic patient is apparently less able to mobilize vivid concept residuals and becomes the victim of contemporary background anchors. A protected hospital setting might be considered the analogue of the constant method, while the more complex, responsible details of everyday living can be compared with the method of limits. Most severely disturbed patients can function within a simple, well-defined context, but show considerable disorder in the face of less structured and more demanding problems.

Schizophrenic patients additionally simplify their behavior by using fewer of the available judgment categories. However, healthy Ss differed from physically disabled and other psychiatric patients as well as from the schizophrenic population. It may be that this finding is as much a function of the severity of an illness and the presence or absence of psychopathology as it is of schizophrenia.

This study suggests that contextual influences upon patient-control judgment differences are not sources of experimental error, but may represent the essence of these differences. Without a representative sample of experimental circumstances and additional control groups, it is presumptuous to generalize to an entire dimension of judgment and a specific illness.

#### E. SUMMARY

The previous finding of greater overestimation of short auditory durations in a method of limits situation by schizophrenic Ss was reexamined by using the fixed, simpler method of constant stimuli. Groups of twenty Healthy, Schizophrenic, Other-Psychiatric, and Physically Disabled Ss judged a randomized series of seven different durations along a nine category scale from *very much less* to *very much more* than a clock second. The following results were obtained:

1. No significant effect due to *Populations* was obtained that would confirm a general tendency for schizophrenic Ss to overestimate short auditory durations, verifying the predicted relevance of psychophysical method in obtaining patient-control differences.
2. Slope differences indicate that schizophrenic Ss use fewer judgment categories than do healthy controls, but further analysis suggests that this may be as much a function of severity of illness and the presence of psychopathology as it is of schizophrenia.

These results are discussed in terms of the implications of contextual and population factors in the study of psychopathology.



## REFERENCES

1. HELSON, H. Adaptation-level as a frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1-29.
2. LHAMON, W. T., & GOLDSTONE, S. The time sense: Estimation of one second durations by schizophrenic patients. *Arch. Neurol. & Psychiat.*, 1956, **76**, 625-629.
3. WEINSTEIN, A. D., GOLDSTONE, S. & BOARDMAN, W. K. The effect of recent and remote frames of reference on temporal judgments of schizophrenic patients. *J. Abnorm. & Soc. Psychol.*, 1958, **57**, 241-244.
4. WRIGHT, D. J., GOLDSTONE, S., & BOARDMAN, W. K. Time judgment and schizophrenia: Step interval as a relevant contextual factor. *J. Psychol.*, 1962, **54**, 33-38.

Moccasin Bend Psychiatric  
Hospital  
Chattanooga  
Tennessee

Baylor University College of Medicine  
Houston State Psychiatric Institute  
1300 Moursund Avenue  
Houston, Texas

## PSYCHOLOGICAL ASSESSMENT OF DIFFERENTIAL IMPAIRMENT IN CEREBRAL ORGANIC CONDITIONS AND IN SCHIZOPHRENICS\*

*Department of Psychiatry, Cornell University Medical College  
and Hunter College*

---

CHARLES A. KNEHR<sup>1</sup>

---

### A. INTRODUCTION

The often repeated experience of being unable to document clinical impressions of psychological impairment by quantitative techniques led to the present series of studies. The difficulty of translating clinical statements about psychopathology into manageable psychological constructs may account for part of the problem, but the research literature also displays ample evidence of contradictory observations relative to what are presumed to be the same phenomena. Such inconsistencies often present a challenge, and as such, were the immediate stimulus to present work.

The initial interest was directed toward achieving a clearer delineation of the nature of cognitive or thinking impairment in acute schizophrenic illness. These data (11) showed some inefficiency on tasks heavily loaded on *g* functions, with no deficits evident in any other tested abilities. The generally negative results were interpreted to mean that cognitive abilities were essentially intact in these patients and to support the inference that somatic factors may be of only incidental importance. The tests were held to be sensitive to their purpose of detecting increased nervous system limitations, the normal performance levels indicating no such alterations.

Early in this work, an occasional patient with clinically diagnosed cerebral pathology was tested with the measures then in use with the schizophrenic population. Their performances were often well outside the range for schizophrenics and controls which had been studied up to that time. In advance of completing data collection on the schizophrenic group, it was conjectured that persons with cerebral damage or dysfunction could be differentiated on the basis of the type of functional impairment. The study was expanded immediately to include cases with clinical evidence for cerebral histopathological changes.

---

\* Received in the Editorial Office on April 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The assistance of Miss Myrtle Guy, Miss Ethel Schwartz, and Dr. Natalie L. Brown is gratefully acknowledged.



Based upon these preliminary observations and several related studies of more limited scope (3, 9, and 10), it was hypothesized that impairment in psychopathological states involves altered organization of otherwise intact cerebral mechanisms (e.g., altered set, attentional or motivational functions) whereas deficits in the brain-damaged result from increased limitations on function.

As stated, the present study is concerned with only one segment of a much larger general problem with ramifications as broad as the field of psychological inquiry. For a detailed account of previous work with especial reference to cerebral pathology and ablations, a quick access to the extensive literature may be obtained through such articles as those by Babcock (1), Klebanoff et al. (8), Yates (23), Meyer (15), Chapman and Wolff (4), Payne and Inglis (18), and Reitan (20).

The path to the measurement of impairment is often strewn with obstacles. Yet, perhaps the single most frequent clinical activity expected of psychologists consists in the assessment of anomalous behavior tendencies and relative deficiencies of individual capabilities. Successful psychodiagnostic evaluation implies an adequate knowledge of the nature, range and interrelations of psychological processes and objectively determined grounds for decision as to when a deviation is abnormal or represents impairment.

There are some circumstances in which such erudition is not needed. If a person is available in his "normal" state in advance of an event known to produce losses, the impairment of psychological functions is not difficult to measure although the procedures may be tedious. The essential requirement is baselines of optimal performance on tests of adequate reliability to serve as standards of comparison. This can be achieved for a variety of psychological functions in advance of conditions responsible for reduced efficiency or deficit by sufficient practice to remove the last traces of learned improvement and to achieve limited trial to trial variation. Once this has been accomplished, the difference between the established base performances and the data of retest will ordinarily provide the required estimate of functional loss or inefficiency. This procedure has seen extensive use in both human and animal studies.

In many clinical situations, however, the investigator appears to suffer from a reduction in degrees of experimental freedom. In areas which call for experimental exploration, satisfactory test data obtained prior to the onset of conditions productive of psychological impairment are seldom available, and it is too late to achieve direct measurement of preimpairment functional efficiency.

Impairment is a relational concept and implies a decrement in or worsening of psychological functioning from some former level, whether the earlier psychological efficiency was low or high. The problem has been viewed as requiring some representation of either the earlier individual "norm" or surrogate base line or some external standard against which a performance may be compared. Since the work of Babcock (1) there has been a rather extensive search for reliable postimpairment performances which can be used as individual surrogate base lines. The assumption has been that the presence or absence of impairment cannot be determined by means of the standing of individual performances relative to that of others, but can only be determined by reference to an individual's own norms. This inference has set the pattern for many typical experimental designs in which the estimate of the person's preimpairment functional efficiency is based upon those individual processes which are presumed to be less sensitive to particular impairment producing situations. Perhaps too little attention has been paid to measures other than intra-individual differential test performances.

In addition, some of the past difficulties in achieving a satisfactory distinction between impairment related to psychopathology and that related to cerebral dysfunction in humans may result from adherence to the idea that the best of all approaches to impairment study is to measure the most complex aspects of mental activity. These "higher order" functions are presumably the most sensitive to cerebral pathology. However, if the hypothesis stated earlier is correct, nervous system limitations should be evident not only in more complex mediated activities, but also in simpler basic functions as well. In schizophrenic patients without histopathological or neurophysiological changes, impairment would be most evident in complex mental functioning, showing up as impaired intellectual efficiency. If differentiation is to be made between psychopathologically and organically induced impairment, then observed impairment of complex functions would need to be supplemented by performances which allow simpler<sup>2</sup> and more basic functions to be measured. These statements bypass complicating factors which may involve afferent and efferent pathways or subcortical damage which may produce effects of unusual configuration. Ignoring these complexities is perhaps justified in attempting to get a toe hold on the nature of psychological impairment observable under different circumstances.

The test battery already in use with schizophrenics appeared to be made up of desirable tests of both types of processes and, a priori, appeared likely

<sup>2</sup> It is probable that none of the functions referred to are simple, but rather require less mediational complexity.



to meet the requirements for a test of the hypothesis that simpler psychological functions, subject chiefly to system limitations, would provide more adequate separation of organics from schizophrenics than tests of "higher" mental ability.

## B. METHOD

### 1. *Subjects*

The organic group was made up of the first 50 patients on whom essentially complete test data were achieved, and for whom there was clinical evidence of cerebral histopathological changes. Presumably, no two of these patients exhibited the same configuration of cerebral insult. The cases ranged from a very successful prefrontal lobotomy and a temporal lobectomy to a number of patients with diffuse damage related to alcoholism, arteriosclerosis and cerebral disease. Only one had suffered traumatic injury. Also included were several presenile dementias (Pick's and Alzheimer's diseases). The mean age of this group was 47.2 years, with a sigma of 11.4.

The control group (the older half of a dual control group) consisted of 25 individuals with an average age of 46.4 years, with a sigma of 12.2. Because it was probable that the vocabulary score itself had been affected in many of the organics, it was not possible to match the controls with the patients except by age, sex, and in a crude way, by preillness indications of general level of mental ability. From our experience with the tests, there is no assurance that close matching is of special benefit except when comparisons are being made relative to general intellectual ability. The best that can be said is that patients in the organic group probably varied from average to very superior in preillness ability and included a number of business and professional people.

Group results for the schizophrenics ( $N = 50$ ) have already been reported in comparison with the younger half of the dual control group. Additional details on these groups may be found elsewhere (11).

### 2. *Tests*

The test battery has been previously described in detail (op. cit.). The different tests are listed here for convenience.

1. Progressive Matrices (1938)
2. Wide Range Vocabulary Test, Form B
3. Flicker Fusion Test Set (after Landis)
4. Oscillation (using Pursuit Rotor)
5. After-Image Test

6. Series Choice Reaction Time
7. Span of Apprehension Test
8. Divided Attention Test
9. Persistence Test
10. Opposites Test

Aside from the levels of test performances themselves, variability on several of these measures was of initial interest and became part of the data intended for analysis.

The Minnesota Multiphasic Personality Inventory (MMPI) was given each individual at the end of the test series chiefly as a screening device for controls, and for what information it could provide on the patients of the experimental group.

### 3. Procedure

The tests were given in the order listed above to both patients and controls. The Progressive Matrices and Vocabulary were generally given to patients in the morning. After a break for lunch, the remaining eight tests were administered in two sessions of about forty minutes each with a rest between Tests 6 and 7. The controls occasionally desired to work without the rest periods. Because of the variety of tasks and the improbability of significant fatigue effects, the omission of rest periods was permitted.

## C. RESULTS

### 1. The Groups

Data for the control and organic groups and the significance of differences between the groups are shown in Table 1.

As might be expected, the patient group with cerebral organic damage was significantly lower in the ability to deduce new relationships and reason by analogy therefrom when compared to controls. But, because of difficulty in matching the groups in terms of earlier intellectual capacity, the relative contribution of organic factors and selection factors to the observed difference in Matrices scores might be questioned. There are clues which indicate that organic factors play the primary role. The mean vocabulary score for the organic group probably reflects some impairment, yet as it stands, it corresponds to a mental age equivalent of  $17\frac{1}{2}$  years. When not taking age into account, the Matrices mean performance corresponds to a mental age of about 11 years. Or, working from interpolations of the age norms, the organic group as a whole shows an IQ of about 95 on the Matrices against an IQ of about 117 (uncorrected for age) as estimated from the mean vocabulary score.



TABLE 1  
THE MEANS, STANDARD DEVIATIONS, DIFFERENCES BETWEEN THE MEANS AND  
THEIR SIGNIFICANCE FOR THE CONTROL AND ORGANIC GROUPS

Test	Controls <i>N</i> = 25		Organics <i>N</i> = 50		Difference	<i>P</i>
	Mean	$\sigma$	Mean	$\sigma$		
Progressive Matrices raw score	50.88	7.87	30.28	9.93	20.60	.001
Wide Range Vocabulary raw score	89.60	5.68	70.56	15.52	19.04	.001
CFF in cycles/second	41.55	3.46	35.95	6.61	5.60	.001
Variability in CFF threshold	.90	.41	.96	.50	.06	—
After-image duration in seconds	29.21*	23.08*	39.63**	24.12**	10.42	—
Series choice reaction time in seconds	.78	.063	1.04	.197	.26	.001
Variability of series choice reaction time	.14	.045	.22	.089	.08	.001
Pursuit learning—mean of Trials 1-5 in seconds	4.95	1.68	2.77	1.77	2.18	.001
Pursuit learning—mean of Trials 16-20 in seconds	8.43	2.03	4.60	2.68	3.83	.001
Variability of pursuit learning	.71	.20	.57	.20	.14	.05
Span of apprehension response error—Part I	.70	.73	.56	1.28	.14	—
Span of apprehension response error—Part II	—0.06	.82	—2.22	1.12	.1	—
Divided attention loss in per cent	33.97	22.78	62.41	38.40	28.44	.001
Opposites—raw score	65.88	9.37	42.04	16.12	23.84	.001
Persistence—minutes	4.50	2.48	3.20	2.38	1.30	.05

\* *N* = 20 Test object not seen within two minutes by 5 subjects in control group.

\*\* *N* = 16 Test object not seen within two minutes by 34 patients in organic group.

When the Progressive Matrices subtest scores are shown separately (figures have been rounded) for the five sets of twelve test items each, marked differences between control and organic performances are evident. The complex discrimination items of Set A were handled fairly well by the

	A	B	C	D	E
Controls	11.7	11.0	10.2	10.0	7.9
Organics	10.2	6.6	6.0	5.4	2.1

organic group. Adults of average intellectual ability would be expected to score 11 or 12 on Set A. In Series B which requires a shift to easy reasoning type items, there was a sharp drop in performance by the organics with the differences between the groups tending to increase as reasoning problems became more difficult. As noted above, even if vocabulary performance is subject to some impairment, the group mean suggests the organic group to be at least above average in preimpairment verbal ability and presumably general intellectual ability. It would seem therefore that the mean difference in Matrices performance represents a very real loss in capacity for the perception of new relationships and abstract reasoning.

The Vocabulary was significantly lower in the organic group and, here again, the inability to find a valid means of matching the groups provides the possibility that some of the observed difference may have resulted from the selection procedure. Vocabulary is subject to the effects of varied training and experience. It is also known to possess an above average resistance to loss producing situations because it is usually a well habituated activity. Present data do not provide a basis for an estimate of the relative contribution of these factors.

The three performances which were included a priori as measures of perseveration, i.e., the CFF, the After-Image duration, and Series Choice reaction time, provided easy separation of the groups. The data for After-Image duration is misleading as shown. The means and sigmas are for the 20 control subjects and 16 patients of the organic group who did see the target. It was discovered that when an individual had not seen the target in the flashing weak monochromatic red light within two minutes, he would not see it at all. These persons apparently never reached the threshold needed to report the target opening. Of those who saw the target, the controls reported it sooner although the difference between the groups was not significant. But it is important to note that 68 per cent of organics failed to respond in two minutes compared to 20 per cent of the controls.

Pursuit learning was markedly different between the control and organic



groups. The patients with cerebral organic conditions found it more difficult to follow the target with the stylus on the initial trials with only a small increase in proficiency during the remainder of the twenty trials, never reaching the mean of the first five trials of the control group.

The group results on the Span of Apprehension were so variable that, for all practical purposes, there were no group differences. Many approaches to the analysis of the data were tried, from correlations between responses and stimuli to the use of direct deviations as shown. There was no consistent quantitative difference in mode of response observed between the two groups.

Marked group differences were observed in the reduction of performance when two tests, which had been performed separately, were then performed simultaneously (Divided Attention). Individual differences were rather large for both groups on this test.

The ability to give Opposites quickly to a set of easy to difficult stimulus words showed a significant decrement in the organic group when compared to the controls.

Persistence in a task which becomes fatiguing and annoying was significantly less in the patient group, but only at the  $P = .05$  level.

The measures of variability evidenced no consistent overall trend. There was no difference in the variability of the CFF threshold, a significant increase in variability of Series Choice reaction time along with a decrease in variability of Pursuit learning. The latter decrease is undoubtedly related to the fact that performance was so minimal that sizeable variability was restricted.

The data for the schizophrenic group and its controls have been previously reported. In order to show the chief findings for normals, schizophrenics and organics together with least confusion, the group means for six tests (excluding the Vocabulary) which yielded highly significant differences ( $P = .001$ ) between organics and their controls are shown in Table 2. Two measures based on the Matrices performance are also shown in addition to the raw score.

Acute schizophrenic patients are generally younger than persons with the varieties of cerebral pathology found in the organic group. In the present instance the difference was 22 years. The schizophrenics were not significantly different from their own controls and did not differ from the older controls for organics except for a higher CFF threshold which appears to be somewhat sensitive to age differences, and a lower Opposites score, both significant.

It is evident from Table 2 that, as a group, schizophrenics do not show the losses characteristic of the organic group, except possibly on Opposites.

TABLE 2  
GROUP MEANS FOR SCHIZOPHRENICS, ORGANICS AND THEIR CONTROLS ON THE INDICATED TEST MEASURES  
(All control-organic and schizophrenic-organic differences are significant.)

Test	Group means		
	Controls	Schizophrenics	Organics
Progressive Matrices:			
Raw score	50.88	49.24	30.28
B-drop (differential between Sets A & B)	.76	.92	3.60
Set B—raw score	11.00	10.50	6.60
CFF in cycles/second	41.55	44.60	35.95
Series choice reaction time in seconds	.78	.76	1.04
Pursuit learning:			
Mean of Trials 1-5 in seconds	4.95	5.07	2.77
Mean of Trials 16-20 in seconds	8.43	8.86	4.60
Divided attention loss in per cent	33.97	22.23	62.41
Opposites—raw score	65.88	53.37	42.04



Those measures which yielded significant differences between the controls and organics of Table 1 distinguished between schizophrenics and organics at equivalent levels of statistical significance.

## 2. *The Individuals*

Highly significant differences have been observed between means of the control and organic groups on a majority of the tests used. A significant difference was observed ( $P = .05$ ) for another. While such results may be gratifying, the ability of such procedures to sort out cerebral histopathological and neurophysiological changes from normals requires reference to individual performances.

The data showing separate test capabilities of classifying individual controls and organics into the correct criterion groups are condensed in Table 3 and are shown as per cent misclassifications relative to the indicated cutting scores.

Taken individually, the tests ranged from 20 to nearly 50 per cent misclassification of organics and yielded from 0 to 12 per cent false positives with the controls. If one were limited to only one test, the Matrices raw score, the Matrices Set B raw score, or the Opposites would seem to hold the greatest promise of separating the organics from the controls. These scores depend to a large extent on the probability that the preimpairment intellectual abilities of these persons was average or above. Unless there were a fair degree of assurance that an individual could be thus classified, these cutting scores could not be used with confidence.

The Series Choice time seems to be nearly as good as the above measures in distinguishing organic impairment. First, a score of .90 second or above did not trap any controls irrespective of age. Second, from tables of intercorrelations for the different groups studied,  $r$ 's between intellectual ability and Series Choice time varied from .01 to .34. Since there is, at most, only a slight relationship, one need not be so careful of preimpairment ability. In taking these factors into account and noting the lower discriminative capabilities of the remaining measures, this test would seem to be the best single measure in the battery from the standpoint of its potential for giving a low misclassification figure for a broader range of preimpairment individual abilities.

The question arises in Table 3 as to whether the control misclassifications represent few or several individuals. Table 4 shows that, while controls occasionally fall on the abnormal side of cutting scores on the tests, no one control subject did so for more than one test. If a criterion of only two

TABLE 3  
PER CENT MISCLASSIFICATION OF ORGANIC AND CONTROL INDIVIDUALS BY SEPARATE  
TESTS WHEN BASED ON INDICATED OPTIMAL CUTTING SCORES\*

Test	Per cent organics misclassified as normal	Cutting score	Per cent controls misclassified as organics
Progressive Matrices:			
Raw score	20	40	4
B-drop (differential between Sets A & B)	44	4	4
Set B—raw score	26	8	4
CFF in cycles/second	48	37	4
Series choice reaction time in seconds	30	0.90	0
Pursuit learning:			
Mean of Trials 1-5 in seconds	38	3	8
Mean of Trials 16-20 in seconds	34	6	12
Divided attention loss in per cent	36	60	4
Opposites—raw score	28	50	4

\* These data are derived from distributions of individual scores on the tests yielding the most significant group differences.



abnormal scores is taken, no controls would be called organic, and four per cent of the organics would escape the net. If a more stringent cut-off point of three test failures is taken, only one in ten of the organics would be missed while increasing the probability that no normal individual would be called organic.

Because different patients failed different combinations of tests, and since the number of their performances on the abnormal side of the cutting scores ranged all the way from one to eight, it would appear that all eight measures would be needed to achieve a 90 per cent discrimination of organic features at severity levels which were used in this study.

If one examines the procedures being followed, it is apparent that the individual data are compared with some empirically determined range of normal function without special reference to estimated individual preimpairment abilities. The score differential on the Matrices (B-drop) can either be thought of as a deviation from some normal range or as the loss of flexible functioning relative to a more resistant function within the same person—possibly discriminative ability.

The individual data for the Matrices and Vocabulary were used to calculate an often-used type of differential in which the individual is compared with himself. Vocabulary, widely held to be relatively resistant to impairing factors, is used to approximate best levels of development, while the Matrices can be taken to indicate the present level of flexible abstract ability. A significant lowering of the Matrices performance relative to Vocabulary might be taken as an indicator of impairment.

IQ's were calculated from vocabulary age and by conversion of Matrices centile data which already takes into account the effects of normal ageing. The net result of this has been to shrink the differences found since it seemed inadvisable to combine data from above average levels, which had to be derived from distributions with below average data which could have been calculated by mean age data down to six years.

The organic group averaged a 17.7 point IQ loss which was significantly greater than the control group loss of 5.12 points. Individually, the losses can be compared as follows:

	Matrices loss in points of IQ relative to vocabulary score			
	0-9	10-19	20-29	30 and over
Controls—per cent	76	12	4	8
Organics—per cent	32	22	22	24

TABLE 4  
 BASED UPON THE SELECTED CUTTING SCORES,\* THE PER CENT DISTRIBUTION OF CONTROLS AND ORGANICS BY THE NUMBER OF  
 TEST MEASURES ON WHICH INDIVIDUALS ACHIEVED SCORES IN THE ABNORMAL DIRECTION

	Number of abnormal scores								
	0	1	2	3	4	5	6	7	8
Controls—per cent	68	32	0	0	0	0	0	0	0
Organics—per cent	0	4	6	6	14	28	24	6	12

\* The Matrices Set B raw score is not included since it is closely related to the Matrices B-drop.



If the cutting score is taken at 19 so as not to misclassify too many controls, the Matrices-Vocabulary differential alone is poorer than any other single measure derived from the test battery. As noted before, there was qualitative evidence that vocabulary was affected to varying degrees in the individuals of the organic group which would tend to minimize such a differential.

Variations in the relative effectiveness of these several test measures to differentiate between organics and controls, makes similar data for the schizophrenics of special interest. As previously indicated, the schizophrenics were much younger than the organics and were not significantly different from the younger control group on any test. There is no way around the age difference. Reference to the data of Table 2 suggests that it is not unfair to judge the individual performances of schizophrenics by cutting scores derived for the older control and organic groups although there were two sizable differences, as noted above, between schizophrenics and older controls with whom they were not matched.

Accordingly, the individual data for all three groups are shown in Table 5 as the per cent of individuals who obtained abnormal scores on the measures indicated. The three scores derived from the Matrices performance show a considerably larger number of scores below the cutting score among the schizophrenics (false positives) than observed for controls. When compared to organics, the percentage of misclassification is relatively small, since less than 20 per cent of schizophrenics respond to the Matrices problems in a manner characteristic of the large majority of organics.

Few controls or schizophrenics were observed to perform at an abnormal level on the CFF, Series Choice and Divided Attention, while 52 to 70 per cent of the organics gave impaired performances. It is evident that an individual found on the abnormal side of the cutting scores for these three tests is likely to evidence some cerebral dysfunction or damage. The schizophrenics' own control group shows misclassification of 4, 0, and 8 per cent respectively on these same three tests, suggesting that age is not a special factor. It would seem that if there were complicating organic factors in a schizophrenic patient, there would be a fairly good chance of their detection with these tests.

The per cent misclassification on the two Pursuit Learning measures seems excessive for both the controls and schizophrenic patients. Even though about two-thirds of the organics make abnormal scores, this test might be considered as the least adequate of this group of tests.

While organics were discriminated from controls fairly well by the Opposites test, it yielded the poorest discrimination between individual schizo-

TABLE 5  
PER CENT OF CONTROLS, SCHIZOPHRENICS AND ORGANICS ACHIEVING ABNORMAL SCORES BY SEPARATE TESTS WHEN BASED ON INDICATED OPTIMAL CUTTING SCORES\*

Test	Per cent controls misclassified as organics	Per cent schizophrenics misclassified as organics	Per cent organics correctly classified
Progressive Matrices:			
Raw score	4	18	80
B-drop (differential between Sets A & B)	4	14	56
Set B—raw score	4	16	74
CFF in cycles/second	4	4	52
Series choice reaction time in seconds	0	2	70
Pursuit learning:			
Mean of Trials 1-5 in seconds	8	24	62
Mean of Trials 16-20 in seconds	12	14	66
Divided attention loss in per cent	4	2	64
Opposites—raw score	4	46	72

\* These data are derived from distributions of individual scores on the tests yielding the most significant group differences.



phrenics and organics. About one-third of all schizophrenics were found within five points of the cutting score of 50. The set for opposites association seems to be fairly sensitive to both functional factors as well as to cerebral dysfunction, although the influence of the latter is more evident.

The distribution of the absolute numbers of tests on which individual schizophrenics achieved abnormal scores may be seen in Table 6. Table 4 shows that no control individuals scored abnormally low on more than one test, whereas 32 per cent of the schizophrenic patients did so. Nearly half of these patients achieved an Opposites score below the cutting score, which accounts for more than a third of these abnormal scores. If the Opposites were dropped from consideration when using the absolute number of test failures to distinguish organics from schizophrenics, there would be no entry under four scores, and the number of schizophrenics showing no abnormal scores would increase to 60 per cent and more nearly approximate the control percentage. Dropping this measure for the organics does not materially increase the frequencies of fewer than four abnormal scores in this group. It is apparent that, except for Opposites, the individual schizophrenics approximate the performance of control subjects much more closely than that of patients with cerebral organic changes.

The Matrices as an IQ score averaged 6.2 points lower than the Vocabulary IQ equivalent for the schizophrenics, which may be compared with a mean lowering of 5.1 points for the controls and 17.7 points for the organics. Using the cutting score obtained from organic-control comparisons, 26 per cent of schizophrenics showed excessive inefficiency on the Matrices compared to 10 per cent of the control group.

#### D. DISCUSSION

As in many other areas of inquiry, experimental study of impairment raises problems in the selection of adequate controls, the specification of characteristics of an experimental population and the size of experimental *N*'s. Accordingly, the adequacy of control procedures may merit some consideration.

There is no doubt that judicious matching of control and experimental subjects on one or several variables is both desirable and necessary for many experimental purposes. It is often questionable, at the same time, whether the desired control has been achieved by every reasonable effort to provide such control matching. It is well known that in persons of roughly equivalent ability, there is variation in original cyto-architecture and initial biases upon which is superimposed differing experiential history. Specific abilities within

TABLE 6  
 BASED UPON THE SELECTED CUTTING SCORES,\* THE PER CENT DISTRIBUTION OF SCHIZOPHRENICS AND ORGANICS BY THE  
 NUMBER OF TEST MEASURES ON WHICH INDIVIDUALS ACHIEVED  
 SCORES IN THE ABNORMAL DIRECTION

		Number of abnormal scores							
		0	1	2	3	4	5	6	8
Schizophrenics—per cent	40	28	14	10	8	0	0	0	0
Organics—per cent	0	4	6	6	14	28	24	6	12

\* The Matrices Set B raw score is not included since it is closely related to the Matrices B-drop.



the matching variables are often dissimilar in pattern and achievement. Attitudes and motivation vary considerably. In addition, if the search for increasingly pure measures of independent dimensions of human ability are realized, controlling one of these, when possible, would have minimal effect on the equalization of the others. If it is assumed that adequate general control can be obtained by the matching of several psychological variables, high correlations between mental functions would be implied, an assumption which runs counter to factorial study.

In patients with clinical evidence for cerebral pathology, test results show impairment in varying degrees and one may not be sure of matching any important variables unless the amount of impairment can be stated, the study of which was the reason for matching in the first place. That some tests "hold" is a relative matter and, although they may be useful, the unknown degree to which a "hold" test has been affected makes this approach inadequate. As an alternative, one might use inferences based upon life history information. Try to find a control subject to match a 41-year-old skilled watchmaker who finished high school by achievement, was always "one for the girls," married twice, had seven children, was a corporal in the army, played poker, but never bet on the horses. He is well aware that his brain is not working as efficiently as it did at one time. What normal subject or hospital patient recovering without incident from major surgery could be found to approximate this individual's former abilities, response tendencies and perceptual modes, or to equalize postimpairment individual attitudes and apprehensions? It is only by the grossest approximation that this could be done, and would probably be based on such inferences as that the patient probably had a reasonably normal and healthy brain at one time, was not stupid, and probably exhibited no clinical levels of psychopathology.

It is not that matching is not needed or is undesirable. Rather, it is often likely that in going through the motions of obtaining matched controls, no real equivalence is produced between groups even though there is serious intent to approximate matching as nearly as possible. The research worker in a good primate laboratory is in a much better position to control these factors since a complete history is often available on every subject from the time of birth.

If truly matched control groups are impossible of achievement, and if "hold" test performances are subject to unknown amounts of loss in a given individual, a loosely contrived control population without patent misfits may be used in another manner. They can be used to determine the most

likely normal range of performances on measures which tend ideally to a narrow range of variability. A normative procedure of this kind has been used as the basis for cutting scores in this study.

The selection of experimental subjects is often dictated by availability tempered by other considerations. There is no assurance that one will ever find two patients with the same locus and extent of cerebral pathology. Where this can be approximated, it is known that similar damage yields varying psychological consequences in different individuals (15). The complexity of brain function is such that sufficient damage in one area may affect integrative activity in general (2, 13) or have less effect than one might anticipate (6). By contrast, primate studies have demonstrated area differences in function (22), and Pribram (19) is disposed to divide higher functions into two main categories with brain areas to subserve these functions. In opposition to studies which show some primary dependence of functions on particular cortical regions, Chapman and Wolff (4) found that outside of motility and sensory functions, impairment of higher processes in humans was not significantly related to the site of damage or right-left localization. In non-surgical conditions, locus of damage is difficult to ascertain and autopsy reports available several years post-test are unlikely to give an adequate basis for describing the pathology at the time of test. Mindful of the varying data on the relations between behavior and locus and extent of cerebral pathology, plus other difficulties in adequately knowing the status in living subjects, it may be that an organic group representing a variety of conditions may have as much merit in our present state of knowledge as a group chosen on some other basis. While not a strictly analogous situation, it may be helpful to remember that Binet's original criterion, teachers' ratings of pupils, was short of being ideal, but it was available. This difficulty has since been overcome and now a teacher's rating would not ordinarily be used in preference to the more accurate and dependable test results. Perhaps such an "operation bootstrap" can be executed successfully in the field of impairment study.

In studying psychological deficits (or other psychological phenomena) fairly large  $N$ 's are required for adequate treatment of data by correlation and other statistical techniques. But, if experimental data are to be useful in distinguishing individuals, functional differences of practical merit should be observable ideally in every one of the first five individuals examined. And on replication, the next five should be differentiable and so the third five. No one seems to have achieved such accuracy with impairment measures.

The above discussion is not intended as support for the dismal outlook



on impairment study implied by such authors as Yates (23) and McFie (14). In spite of and with these difficulties, useful and meaningful data can be obtained and advance in understanding these problems can be achieved. It is believed that the present study is among those representative of some progress.

The measures yielding significant group differences appear to be sensitive to the limitations imposed by cerebral pathology. The two tests which appeared to fall short in this respect are undoubtedly more adequate than indicated by the data shown. After the data collection was well started, it became apparent that there were error sources built into the After-Image test which may have accounted for the excessive number of false positives, and perhaps for a number of the true positives. An improved test of visual adaptive functions might prove to be quite useful. Experience has also indicated that, by inspection of Span of Apprehension response plots, more anomalous response patterns were evident in organic patients than were revealed by any quantitative treatment of the data tried. If a more adequate procedure for evaluating responses can be worked out, the usefulness of this test might be increased. Patterning of deviant responses on this test often has an individual stamp, but the implication of these patterns has resisted correlational study.

When the test data, in combination or singly, are viewed in terms of their success in predicting the criterion of cerebral pathology in a patient group with divergent difficulties, considerable variation is noted. By sheer counting of the number of abnormal scores found among nine measures derived from six tests, nine out of 10 patients with cerebral pathology could be distinguished without misclassifying any controls. If one were in a hurry, five minutes of cooperative patients' time spent on the Series Choice reaction time alone would probably indicate organic factors in seven out of 10 without much danger of suggesting damage in an individual with an intact nervous system. Based on this type of data, the CFF would appear to be still less effective by being correlated with a variety of forms of cerebral dysfunction only about half the time. It should be remembered that the percentages of test hits and misses depend upon optimal location of cutting scores on a continuum of test performances.

Such is the manner in which data of this type are often interpreted. But the ready assumption that test sensitivity is indicated by success or failure in predicting members of the criterion group may not be wholly justified (5). There is evidence from other studies to indicate that the psychological consequences of similar ablations or locus and extent of pathology

vary considerably (6, 16, 17). Test "misses" could just as easily be related to individual variations as to test sensitivity. The failure of measures with group significance to show the same uniformity of loss pattern from one patient to another implies some localization of function. In addition, knowledge about the character of dysfunction or pathology or the stability of locus and extent of cerebral lesions cannot be stated with confidence by any known present technique.

In further appraisal, consider the individual data for Series Choice. There were three test misses in 10 for organics. Assuming cooperation, if the speed of self-initiated choice reaction is limited by the capacity of relevant systems of the central nervous system, then any histopathological or neurophysiological changes which would lower that capacity should be reflected in a slower reaction time. But three out of 10 patients responded in the normal range. One may choose between saying that these individuals appeared to work at levels above a reduced capacity, i.e., the test is inadequate, or that the test is sensitive and the finding of normal speed of response reflects the integrity of the pertinent neural mechanisms.

There is little in the data of this study that requires interpretation in terms of test sensitivity in contradistinction to their implication for individual differences. Rather, there may be cogent reasons for the latter approach which assumes full test sensitivity to both large and small variations in capacity. When the measured loss is greater than some empirically determined amount, the psychological function could then be said to be impaired. It may be further determined that, once observed, such deficit is likely to implicate some particular impairing condition. It follows from this line of thought that with tests of high reliability (repeatability) as a requirement, observed variations in test performance from one individual to another relate to functional differences among individuals and do not necessarily refer to the validity of the tests.

If present data are interpreted from this viewpoint, the individual who makes several abnormal scores shows impaired performance in those several functions with the added probability that the impairment is related to cerebral dysfunction. Or, an individual with clinical signs of cerebral pathology is quite likely to show impaired performance on at least two or three of these different measures. If the individual performs in the normal range on one or more tests, then for these functions he is undoubtedly without any cerebral dysfunction which would interfere with efficient performance. Stated inelegantly, if a motivated patient achieves an abnormal score, he is probably impaired, and if he doesn't, there is no sign that he is. There is no



point in representing an unimpaired performance as really impaired while ascribing the negative finding to test inadequacy.

Such considerations as these suggest that the measurement of deficits, as such, is not attended by insurmountable difficulties. It might appear, however, that this viewpoint is not equally applicable to all types of quantifiable psychological activity. Even with those mental processes involving perception of complex relationships, the maintenance of perceptual sets or the recombination of familiar experiences, the more reliable tests are likely to detect small losses of efficiency. Here, however, variability among individuals is often so great that a normative approach may not be adequate and an estimate of the personal norm becomes more important. On the other hand, there must be variations among psychological functions in their immediacy of relationship to the integrity of central nervous mechanisms. Somewhat removed, in this sense, are undoubtedly many well habituated abilities, e.g., vocabulary. In any case, once impairment is observed, the setting in which the impairment occurs becomes important.

The adequacy of the schizophrenic data for use in testing the hypothesis of discoverable differentials is attenuated from the start by the failure to find any type of gross impairment for the group as a whole. The individual data, however, provide some clues in support of the hypothesis.

If the data are used in the manner suggested for information about individuals rather than the tests, it is apparent that there was a noticeably greater frequency of abnormal scores among schizophrenics than normals on tests of *g* and the learning of a pursuit task. At the same time, impaired performances were usually not observed on CFF, Series Choice, and Divided Attention. The organics, on the other hand, showed a high frequency of abnormal scores on all tests.

There is clear suggestion of a differential here even though impairment was found in only a small minority of schizophrenics. The decision as to whether the impairment is in a setting of psychopathology alone or involves complicating cerebral organic factors would depend on the presence or absence of impairment on those tests normally not impaired in the schizophrenic, but showing fairly high incidence in organics.

Viewing these data in a different way, it was observed that the most frequent type of impairment found in individuals with a wide variety of central nervous system pathology was in the ability to educe relations and correlates, a loss shown by 72 to 80 per cent of these patients. Yet it was in precisely these processes that the greatest incidence of loss was observed in individual schizophrenics. One is forced to the conclusion that measurement

of the efficiency of these processes is least likely to distinguish brain damage from functional impairment.

If this inference is correct, it provides a possible reason for the failure of some studies to develop successful procedures for making such distinctions. It has been noted before (12) that the great majority of studies in this area have been based on tests of abstract intellectual ability, after which observation these authors proceeded to report yet another experiment based on conceptual functions.

Within the framework of factorial studies, one would not expect to find some single measure of universal utility in the assessment of impairment. The present evidence is consistent with this view and demonstrates that a unitary view of psychological function is not compatible with the findings. A recent review of the localization vs. mass-action problem may be found in Meyer (15). In introducing their findings which lead to the mass action view, Chapman and Wolff (4) consider these same problems in detail.

The conclusion stated in the previous paper (11) to the effect that organic factors appear to play at most a minimal role in acute schizophrenic illness appears to gain additional confirmation from present evidence. The previously observed Matrices B-drop pattern among organic patients was not observed in as high a percentage of organic patients in the present group (9). This measure continues to suffer from too many impaired performances in psychiatric patients as found with most of the tests of *g* in this study. About the same proportion of organics in the present study showed elevated scores on the Series Choice as reported previously for an independent study (10). The region of overlap with controls remained quite stable from one study to the other. A consistent majority (70 per cent) of organic damage cases of different types have thus shown impairment of self-initiated choice reaction time.

Such replication of results is, of course, comforting. But the observations and interpretations presented here, even if well taken, require further independent study. Differences in constructs are often more representative of diversity in interpretation than marked variations in data. It would seem that the reappraisal made with respect to data of the type often used for asserting the validity of tests will show, on further study, that the measurement of impairment can be reasonably precise. The problem will be complicated somewhat when introducing the dull and the dim-witted and in studying impairment encountered during developmental years.

In light of the above considerations, it is difficult to share the pessimism of McFie (14) who writes, "Few neurologists today would consider that



psychological assessment constitutes an important part of the investigation of a patient with suspected brain damage." Among other things, McFie is disappointed in the failure of psychodiagnostic procedures to provide localizing information. Perhaps worth pondering is Hebb's statement (7, p. 454) that "... the brain like large government may not be able to do simple things in a simple way."

#### E. SUMMARY

Fifty patients with varying kinds of cerebral pathology and a control group of 25 individuals without known damage were examined with a battery of ten tests previously used with a group of schizophrenics.

The battery included tests traditionally held to measure abstract conceptual activity, along with measures of other aspects of cognitive functioning.

Individual performances on tests showing very significant group differences ( $P = .001$ ) between normals and organics of the same age indicated that some impaired psychological functions accompany a variety of organic conditions with varying frequency.

The assumption that the validity of tests used for assessing impairment can be estimated from test "hits" and "misses" for individuals of a criterion group is questioned. It is urged, rather, that tests of high reliability are probably adequately sensitive, and that a test "miss" could easily signify that the measured function is not impaired. Using the latter assumption, test data then estimate the efficiency of measurable psychological functions and do not refer to test sensitivity.

The hypothesis that impairment in schizophrenics differs in some respects from that of organics received some support. Deficits in schizophrenics, when found, occurred primarily with tests of abstract conceptual functions and complex learning, areas tending to greatest loss in organics. Measurement of these functions provided the least probability of separation of the two groups, while differentiation could be made by the presence of abnormal scores on non-*g* tests on which schizophrenics achieved uniformly normal scores, while abnormal scores, when obtained, were found primarily in clinically verified organics.

Impairment accompanying loss of cerebral tissue or neurophysiological changes is apparently not unitary. The measurement of any one aspect of cognitive functioning will probably not suffice in the detection of organic factors in clinical patients. This applies especially to the use of single tests of intellectual ability if distinction between functional and somatic factors is desired.

## REFERENCES

1. BARCOCK, H. An experiment in the measurement of mental deterioration. *Arch. Psychol.*, 1930, **117**, pp. 105.
2. BATTERSBY, W. S., TEUBER, H. L., & BENDER, M. B. Problem solving behavior in men with frontal or occipital brain injuries. *J. Psychol.*, 1953, **35**, 329-351.
3. BROWN, N. L. A study of sensory, motor and ideational perseveration in schizophrenic and diffuse brain damaged patients. Univ. of Penn. Ph.D. Thesis, 1956, Mic. 56-2352, Microfilm Library of Congress, Washington, D.C.
4. CHAPMAN, L. F., & WOLFF, H. W. The cerebral hemispheres and the highest integrative functions of man. *A.M.A. Arch. Neurol.*, 1959, **1**, 357-424.
5. EREL, R. L. Must all tests be valid? *Amer. Psychol.*, 1961, **16**, 640-647.
6. HERR, D. O. Man's frontal lobes. *Arch. Neurol. Psychiat.*, 1945, **54**, 10-24.
7. ———, Alice in Wonderland or Psychology Among the Biological Sciences. In Harlow, H. F., & Woolsey, C. N., (Eds.), *Biological and Biochemical Bases of Behavior*. Madison: Univ. Wisconsin Press, 1958.
8. KLEBANOFF, S., SINGER, J. L., & WILENSKI, H. Psychological consequences of brain lesions and ablations. *Psychol. Bull.*, 1954, **51**, 1-41.
9. KNEHR, C. A. Progressive Matrices findings associated with cerebral histopathology. *Percept. & Mot. Skills*, 1956, **6**, 249-254.
10. KNEHR, C. A., & BROWN, N. L. Perseveration in schizophrenic and brain damaged patients. *J. Psychol.*, 1957, **43**, 269-275.
11. KNEHR, C. A. Thinking impairment in acute schizophrenic illness. *J. Psychol.*, 1962, **53**, 15-28.
12. LEVENTHAL, D. B., MCGAUGHRAN, L. S., & MORAN, L. J. Multivariable analysis of the conceptual behavior of schizophrenic and brain damaged patients. *J. Abn. & Soc. Psychol.*, 1959, **58**, 84-90.
13. LINDSLEY, D. B. Common factors in sensory deprivation, sensory distortion and sensory overload. In *Sensory Deprivation*. Cambridge: Harvard Univ. Press, 1961.
14. MCFIE, J. Psychological testing in clinical neurology. *J. Nerv. & Ment. Dis.*, 1960, **131**, 383-393.
15. MEYER, V. Critique of psychological approaches to brain damage. *J. Ment. Sci.*, 1957, **103**, 80-109.
16. MEYER, V., & YATES, A. J. Intellectual changes following temporal lobectomy for psychomotor epilepsy. *J. Neuro. Neurosurg. Psychiat.*, 1955, **18**, 44-52.
17. NIELSEN, J. M. Agnosia, Apraxia, Aphasia. New York: Hoeber, 1946.
18. PAYNE, R. W., & INGLIS, J. Testing for intellectual impairment—some comments reconsidered. *J. Ment. Sci.*, 1960, **106**, 1134-1138.
19. PRIBRAM, K. H. Neocortical Functions in Behavior. In Harlow, H. F. & Woolsey, C. N. (Eds.), *Biological and Biochemical Bases of Behavior*. Madison: Univ. Wisconsin Press, 1958.
20. REITAN, R. M. Psychological Deficit. *Ann. Rev. Psychol.*, 1962, **13**, 415-444.
21. SPEARMAN, C. The Abilities of Man. New York: Macmillan, 1927.
22. WARREN, J. M., & HARLOW, H. F. Learned discrimination performance by monkeys after prolonged postoperative recovery from large cortical lesions. *J. Comp. Psychol.*, 1952, **45**, 119-126.
23. YATES, A. J. The validity of some psychological tests of brain damage. *Psychol. Bull.*, 1954, **51**, 359-379.

Payne Whitney Psychiatric Clinic  
 The New York Hospital  
 525 East 68th Street  
 New York 21, New York



## PERSONAL ATTRIBUTES OF COLOR AND DESIGN PREFERENCES IN CLOTHING FABRICS\*

*Department of Textiles and Clothing, University of Maryland*

---

NORMA H. COMPTON

---

### A. PURPOSE

The existence of a relationship between color and design preferences and personality appears to be generally accepted by psychologists, social scientists, and fashion designers. However, the problem as to the specific nature of this relationship has been relatively untouched by scientific investigation. In an attempt to indicate a direction for future studies in this field, the instant study was undertaken in 1961 and the results, while not conclusive, evidence sufficient promise to warrant further investigation. The primary purpose of this exploratory study was to determine whether or not relations exist between preferences for certain aspects of color and design in clothing fabrics and selected physical and personality characteristics of the individual.

### B. METHOD

Participants in this study were 145 freshmen women in the College of Home Economics at the University of Maryland.

An instrument consisting of paired fabric choices was designed to determine students' fabric preferences. Three series of fabrics, 5" x 7" in size, were mounted on 8½" x 11" cards. One group of choices involved selection of plain weave cotton fabrics of the six hues of red, yellow, orange, blue, green, and purple, with respect to a tint, shade, or saturated hue. For a given hue, each of the three fabrics was presented for comparison with each of the other two fabrics, so that a total of eighteen paired comparisons was made. Another series consisted of fifteen choices of fabric designs with strong contrasts in the value of color between the background and the pattern as opposed to the same designs with weak value contrasts. A third series consisted of fifteen choices between large and small patterns of striking similarity. Each fabric of a pair within a series and cards for all series were randomized before presentation to students for choices.

Students were classified into preference groups only if they exhibited

---

\* Received in the Editorial Office on April 30, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

very definite and consistent preferences on two spaced administrations of the fabric preference instrument. To determine the score (or number of choices) which would constitute a strong preference in each series of choices, the binomial expansion method of probability determination was used. At a selected probability level of .057, a score of ten or more points or choices out of a possible twelve preference points for tints, shades, or saturated hues was deemed necessary for inclusion in one of these preference groups. At a probability level of .036, a score of twelve or more choices out of fifteen possible choices for strong or weak color value contrasts and for large and small designs was required for classification into any of these groups. These scores had to be maintained on the second administration of the instrument, which followed after a two weeks' interval. Test-retest reliability coefficients, calculated for each classification from a random sample of students, ranged from .81 to .92.

After the students were grouped according to strong color and design preferences, these preference groups were compared to determine if they also differed in eye color, hair color, and weight/stature ratios; in their interest in the occupational fields of clothing merchandising, designing, and textile testing; and in the personality dimensions of dominance, sociability, self-acceptance, good impression, flexibility, and femininity. Occupational interests were determined from students' scores on the Johnson Home Economics Interest Inventory. Personality measures were selected from the California Psychological Inventory.

Chi-square contingency tables were set up to test each of the color and design variables for independence with respect to each physical characteristic. The occupational interest scores and the personality measures for the three color preference groups of Tint, Shade, and Saturated hue were analyzed for differences in mean values by analysis of variance. T-tests were made to analyze the difference between mean values of the scores of the Strong and Weak color value contrast groups and between the scores of the Large and Small design groups.

### C. RESULTS AND DISCUSSION

At the five per cent level of confidence, preferences for all of the color and design variables under investigation were independent of all of the physical characteristics (eye color, hair color, and weight/stature). Such physical factors are sometimes assumed to influence the selection of clothing fabrics, although there appear to be no research findings at the present time to support this assumption.



In spite of the generally acknowledged limitations of personality and interest measurement and the questionable correlation between inner personality and outward behavior, several statistically significant differences were found in personality and occupational interests between certain color and design preference groups. Students in the Small design size preference group scored significantly higher in interest in clothing merchandising than students in the Large design size preference group ( $p < .05$ ). It may be that this aspect of the fabric preference instrument was tapping some personality variable in terms of which these students were alike or there may exist some higher order determinant of both design size taste and this occupational interest, with both the lower order variables having found expression in the scores of these students. Table 1 is a presentation of the mean values and variances of clothing merchandising interest scores for students in the Large and Small design size preference groups.

TABLE 1  
FABRIC DESIGN SIZE PREFERENCES OF FRESHMEN HOME ECONOMICS STUDENTS  
IN RELATION TO THEIR SCORES IN CLOTHING MERCHANDISING

Johnson inventory measure	Fabric design size preference group	
	Large	Small
Clothing Merchandising:		
Mean	300.13	318.50
Variance	214.70	252.29

Students preferring deep shades and saturated colors scored higher in sociability than those preferring tints ( $p < .01$ ). People with high sociability scores were described in the California Psychological Inventory Manual as being outgoing, forward, and sociable. Low scorers tend to be seen as conventional, quiet, submissive, detached, and passive in attitude. Therefore, these color choices would appear to be an outward expression of this inner personality trait. A summary of the analysis of variance test is given in Table 2.

TABLE 2  
ANALYSIS OF VARIANCE FOR EQUALITY OF MEAN SOCIABILITY SCORES OF  
COLOR PREFERENCE GROUPS

Sources of variation	Degrees of freedom	Sum of squares	Mean square	F
Between Groups	2	1,109.47	554.74	11.23
Within Groups	33	1,629.51	49.38	
Total	35	2,738.98		

Students preferring small designs scored significantly higher on the good impression personality measure than students preferring large designs ( $p < .05$ ). Inclusion of this personality scale in the California Psychological Inventory developed from a study concerned with identifying responses which adolescents characteristically employ when attempting to present themselves in the best possible light. Among the factors which emerged was a desire to present the self as unaffected, natural, and modest. This characteristic appears to have been reflected in a preference for smaller, less bold designs by students desiring to make a good impression than by students scoring low in this dimension. This result appears to lend support to the concept of the role played by clothing items in helping the individual conform to a cultural ideal, thus expressing those traits which seem desirable to others.

A statistically significant difference was found between mean femininity scores of the Large and Small design preference groups ( $p < .01$ ). Students with high femininity scores preferred more small designs than students with low scores in femininity. Table 3 is a presentation of the mean values and variances of California Psychological Inventory Scores on good impression and femininity for students preferring large and small designs on fabrics.

TABLE 3  
FABRIC DESIGN SIZE PREFERENCES OF FRESHMEN HOME ECONOMICS STUDENTS  
IN RELATION TO THEIR SCORES ON TWO MEASURES FROM THE  
CALIFORNIA PSYCHOLOGICAL INVENTORY

C.P.I. measure	Fabric design size preference group	
	Large	Small
<i>Good Impression:</i>		
Mean	14.14	20.13
Variance	22.48	19.27
<i>Femininity:</i>		
Mean	20.43	26.38
Variance	6.29	11.12

Although this study was exploratory in nature and was based upon a small number of participants, the results would appear to contribute some experimental evidence to support the concepts that the self is expressed through one's selection of clothing fabrics and that clothing items play a role in helping the individual to conform to an ideal self. Such results should help to set the direction for much needed future research in this relatively untouched area of investigation. Understandings of the significance of textiles and clothing in the lives of individuals have strong implications



for human development, for the teaching of principles of textiles and clothing selection, and for designers and merchandising personnel who need to analyze public appeal for clothing articles.

## REFERENCES

1. COMPTON, N. H. Clothing Fabric Preferences in Relation to Selected Physical and Personality Characteristics. Unpublished doctor's dissertation, Univ. Maryland, 1962.
2. GOUGH, H. G. C.P.I. Manual. Consulting Psychologists Press, Palo Alto, Calif., 1957.
3. ———. California Psychological Inventory. Consulting Psychologists Press, Palo Alto, Calif., 1956.
4. ———. "On Making a Good Impression." *J. Educ. Res.*, 1952, 46, 33-42.
5. JOHNSON, H. Johnson Home Economics Interest Inventory. Ames, Iowa: Iowa State College Press, 1955.

*Textiles and Clothing Department*  
*University of Maryland*  
*College Park, Maryland*

## SOME FACTORS AFFECTING THE PERCEPTION OF EVENTS AS CHANCE DETERMINED\*<sup>1</sup>

*Department of Psychology, Ohio State University*

---

SHELDON BLACKMAN

---

### A. INTRODUCTION

A series of recent studies has demonstrated that whether or not success on a task is perceived as a function of one's skill, or is seen as a function of chance affects the extinction of expectancies for success. Neff (3) using a skill task, color matching, found no significant differences in trials to extinction of expectancies for success between partially reinforced groups versus continuously reinforced groups. Phares (4) instructed some Ss that success on color matching and length judging tasks was a function of their skill and others that success on these tasks was a matter of luck. He found that when Ss were instructed that success was a matter of chance, changes in verbalized expectancies as to success were smaller and the number of unusual shifts, down after success or up after failure, was more frequent. James (2) and James and Rotter (2) again found that when Ss perceive a task as one in which success is due in some measure to chance, extinction of expectancies is delayed after partial reinforcement and the number of unusual shifts is higher than when success is perceived as a function of skill. Rotter, Liverant and Crowne (6) used two different tasks which were hypothesized as skill and chance tasks on the basis of the Ss' culture rather than differences in instructions. The main hypothesis that under chance conditions there is less learning by experience was supported.

From a logical point of view one might explain these findings by the gambler's fallacy operating in situations in which success is perceived as a function of chance. In chance situations Ss are likely to have rising expectancies that a change is due after one of two possible events has occurred for a large number of times.

---

\* Received in the Editorial Office on May 1, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This article is based on a dissertation presented in partial fulfillment of the requirements for the degree Doctor of Philosophy in the Graduate School of The Ohio State University. The author wishes to acknowledge a special debt of gratitude to Dr. Julian B. Rotter, who expended a great deal of effort directing this study, and to acknowledge the assistance of Professors Robert J. Wherry and Shephard Liverant, who served as members of the advisory committee.



It seems important to know what some of the conditions are which determine the perception of events as chance determined. The conditions studied here were, length of sequences of events, and whether the sequences were patterned. From the research described above it can be concluded that if a condition in training led to a slow rate of extinction, *Ss* tend to perceive that series as more a function of chance than a training series which leads to a faster rate of extinction. Long sequences and patterned sequences seem to be more easily seen as series on which successful guessing is a function of *S*'s skill, at least skill in guessing what the experimenter will do, than are series with short sequences and non-patterned sequences.

*Ss* were presented with varying series of red and green lights in training, then the red light stopped coming on. The major hypothesis states: *Ss* who received long sequences or patterned sequences of red and green lights will, when the red light stops coming on, have a lower expectancy for red than will *Ss* who received short sequences or non-patterned sequences.

### B. METHODOLOGY

In order to test the hypothesis 180 male and female students from an introductory psychology course were used as *Ss*. Two 7- $\frac{1}{2}$  watt bulbs, one red and one green were placed 20 inches apart on a 10- $\frac{1}{2}$  by 24 inch board. Each *S* was instructed to note on a sheet placed in front of him which light he expected to go on. He was also instructed to note how sure he was that he was correct on a scale from zero to ten, ten being most sure. Eight groups of 20 *Ss* each were arranged in a factorial design as described in Table 1.

Number of sequences was introduced as a control variable since in any set number of trials the groups with short sequences would receive more sequences than the groups with long sequences. The pattern was 3, 2, 1, 1, 2, 3, i.e., three green lights followed by two red, etc. In interviewing, this pattern was found to be too difficult and twenty *Ss* were run in an easy pattern group. The easy pattern was 3, 2, 2, 3. The sequence length for this group was short, and the number of sequences was thirty. This easy pattern group is then comparable to Group I.

After the series of red and green lights, each of which ended with the red sequence of maximum length, the green light was turned on sixty times in succession. Three dependent measures were used to test the hypothesis: (1) the number of red responses in extinction, (2) the mean expectancy for success associated with each of the red guesses, and (3) the number of trials to an extinction criterion of eight successive guesses of red.

TABLE 1  
AN OUTLINE OF THE FACTORIAL DESIGN

<i>Variables</i>			
A—Sequence length			
A <sub>1</sub> 1, 2, or 3 in a row			
A <sub>2</sub> 5, 6, or 7 in a row			
B—Number of sequences			
B <sub>1</sub> 30 sequences			
B <sub>2</sub> 10 sequences			
C—Patterning			
C <sub>1</sub> Not patterned			
C <sub>2</sub> Patterned			
<i>Experimental groups</i>			
Group	I	A <sub>1</sub> B <sub>1</sub> C <sub>1</sub>	
Group	II	A <sub>1</sub> B <sub>2</sub> C <sub>1</sub>	
Group	III	A <sub>1</sub> B <sub>1</sub> C <sub>2</sub>	
Group	IV	A <sub>1</sub> B <sub>2</sub> C <sub>2</sub>	
Group	V	A <sub>2</sub> B <sub>1</sub> C <sub>1</sub>	
Group	VI	A <sub>2</sub> B <sub>2</sub> C <sub>1</sub>	
Group	VII	A <sub>2</sub> B <sub>1</sub> C <sub>2</sub>	
Group	VIII	A <sub>2</sub> B <sub>2</sub> C <sub>2</sub>	

### C. RESULTS

Three separate three-way analyses of variance were computed, one for each of the three dependent variables. (See Tables 2, 3, and 4).

The easy pattern group when compared with its analogous non-patterned group, Group I, showed significant differences in the number of red responses in extinction and in the number of trials to the extinction criterion.

TABLE 2  
SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF THE NUMBER OF RED  
RESPONSES IN EXTINCTION

Source	df	Mean square	F
Sequence length (A)	1	483.0	23.9 <sup>b</sup>
Number of sequences (B)	1	93.0	4.6 <sup>a</sup>
Patterning (C)	1	16.9	0.8
A × B	1	24.1	1.2
A × C	1	44.1	2.2
B × C	1	19.7	1.0
A × B × C	1	0.8	0.0
Within cells	152	20.2	
Total	159		

<sup>a</sup> Significant at the five per cent level.

<sup>b</sup> Significant at the one per cent level.



TABLE 3  
SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF THE MEAN EXPECTANCY  
FOR RED RESPONSES IN EXTINCTION

Source	df	Mean square	F
Sequence length (A)	1	47.3	6.7 <sup>a</sup>
Number of sequences (B)	1	35.1	4.9 <sup>a</sup>
Patterning (C)	1	0.5	0.1
A × B	1	0.5	0.1
A × C	1	0.0	0.0
B × C	1	17.6	2.5
A × B × C	1	15.0	2.1
Within cells	152	7.1	
Total	159		

<sup>a</sup> Significant at the five per cent level.

The mean number of reds for Group I is 11.1, the mean for the easy pattern group is 6.6. This yields a *Z* of 3.24 with a Mann-Whitney *U* Test which is significant at the .01 level. The mean number of trials to criterion for Group I is 19.8 and 15.4 for the easy pattern group. A Mann-Whitney *U* Test yields a *Z* of 2.80 which is significant at the .01 level.

The James-Phares Scale (1) and the Internal-External Scale (5) were administered to all *Ss*. These scales were designed to measure how individuals differ in their tendency to view events as chance determined. There were no significant relationships between these scales and behavior in extinction.<sup>2</sup>

#### D. DISCUSSION

Sequence length affects *Ss* in the predicted manner for two of the dependent measures. *Ss* exposed to the long sequences gave significantly fewer

TABLE 4  
SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF THE TRIALS TO CRITERION  
FOR RED RESPONSES IN EXTINCTION

Source	df	Mean square	F
Sequence length (A)	1	0.0	0.0
Number of sequences (B)	1	53.2	1.6
Patterning (C)	1	50.6	1.6
A × B	1	72.6	2.2
A × C	1	76.0	2.3
B × C	1	52.6	1.6
A × B × C	1	19.9	0.6
Within cells	152	32.5	
Total	159		

<sup>2</sup> This research was aided in part by the use of data collected for the United States Air Force Contract No. AF 49 (638)-741 monitored by the AF Office of Scientific Research of the Air Research and Development Command.

red responses in extinction. The expectance associated with the red responses was significantly lower for these Ss. The number of trials to criterion revealed no significant differences. Another more stringent trial to criterion was computed. This was ten red guesses in succession, each one of which was associated with an expectancy of nine or ten. Again there were no significant differences.

A complex pattern yielded no significant differences. However, Ss exposed to an easy pattern gave fewer red responses and fewer trials to criterion than a non-patterned group. In this instance the expectancy for red yielded no significant differences.

#### E. SUMMARY

One hundred eighty Ss were exposed to series of red and green lights which appeared with equal frequency. The series were such that three independent variables were arranged in a factorial design so that their effects could be studied singly and in all combinations. The interest here was on the effect of these variables on extinction. Extinction began when the red light ceased to go on. The dependent measures used were the number of red responses in extinction, the expectancy associated with these red responses, and the number of trials necessary to reach an extinction criterion.

From previous research it was hypothesized that Ss who received long sequences or patterned sequences would evidence a faster extinction of expectancies. These Ss would tend to perceive the experimental task as one in which skill plays a part, as one which is predictable.

It was found that sequence length and the number of sequences significantly affect the number of red responses in extinction and the expectancy associated with them. The more sequences, and the shorter the sequences, the more red responses and the greater the expectancy associated with them. A simple pattern resulted in fewer red responses in extinction and a fewer number of trials to an extinction criterion. The results appear to be understandable on the basis of a skill orientation developing on the part of Ss who received long sequences, or sequences readily perceived as patterned and a chance orientation developing on the part of Ss who received short sequences or non-patterned sequences. Number of sequences was introduced as a control variable and does not seem relevant to the skill-chance dimension.

#### REFERENCES

1. JAMES, W. Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation, Ohio State Univ., 1957.
2. JAMES, W., & ROTTER, J. B. Partial and one hundred per cent reinforcement under chance and skill conditions. *J. Exp. Psychol.*, 1958, **55**, 397-403.



3. NEFF, J. Individual differences in resistance to extinction as a function of generalized expectancy. Unpublished doctoral dissertation, Ohio State Univ., 1954.
4. PHARES, J. Expectancy changes in skill and chance situations. *J. Abnorm. & Soc. Psychol.*, 1957, **54**, 339-342.
5. ROTTER, J. B., SEEMAN, M. R., & LIVERANT, S. Internal versus external control of reinforcements: A major variable in behavior theory. In N. F. Washburne (Ed.), *Decisions, Values and Groups*, Vol. 2. London: Pergamon Press, 1961.
6. ROTTER, J. B., LIVERANT, S., & CROWNE, D. P. The growth and extinction of expectancies in chance controlled and skill tasks. *J. Psychol.*, 1961, **52**, 160-177.

*Mental Hygiene Consultation Service*

*Ireland Army Hospital*

*Fort Knox, Kentucky*

## SEX DIFFERENCES IN RESPONSE SET\*<sup>1</sup>

*Trenton State College*

PHILIP LANGER

### A. PROBLEM

Response set has been generally considered to be an internalized style of test response which is somewhat independent of the stimulus item. Set has been operationally defined to include such response tendencies as agree-disagree, and the acceptance or rejection of items on the basis of their social desirability (1).

While early work stressed primarily the significance of set as a major source of test variance, later results have included relationships between set and specific personality traits or syndromes. Implicit in both lines of research is the assumption that set is a fairly stable phenomenon, at least within the same test or testing situation.

During the course of previous research, a small side project involved an attempt to induce set modification along the agree-disagree dimension. This limited study indicated that this dimension of set could be modified, and further, that females, unlike males, did follow a predictable pattern of change.

It was decided to propose this as a hypothesis, recognizing that small groups and poor controls usually do not lead to significant findings. Our hypothesis, then, is that there are sex differences with regard to the modifiability of the agree-disagree dimension of response set.

### B. METHOD

The Gough Adjective Checklist was administered twice within a month to women ( $N = 95$ ) and men ( $N = 35$ ), enrolled in the general psychology course at Trenton State College. The Checklist consists of 300 alphabetically arranged self-descriptive adjectives, of which 75 have been previously judged socially favorable self-descriptions, and 75 socially unfavorable.<sup>2</sup>

The S first scored the test in accordance with the directions to present an image of "real self." Two weeks later he was asked to present an "ideal

\* Received in the Editorial Office on May 4, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Based on paper given at the April, 1961, EPA.

<sup>2</sup> The reliability of these adjectives is of the order of .90.



self" image. In either case the testee could do so by selecting favorable adjectives, or rejecting unfavorable ones. Thus the content score, which is the measure of self-concept, equals the combined number of *selected* favorable adjectives and *rejected* unfavorable ones.

A number of techniques for measurement of set have been evolved. This study employed an analysis proposed by Fredericksen and Messick (1). The formulae, as revised slightly to meet the specific needs of the study, are:

$$\text{Content} = \frac{\begin{array}{c} \# \text{ of Favorable} \\ \text{Adjectives} \\ \text{(F.A.) accepted} \end{array}}{\begin{array}{c} \text{Total number of F. A.} \\ \text{available} \end{array}} + \frac{\begin{array}{c} \# \text{ of Unfavorable} \\ \text{Adjectives, (U.A.) rejected} - 1 \end{array}}{\begin{array}{c} \text{Total number of U. A.} \\ \text{available} \end{array}}$$

$$\text{Set} = \frac{\begin{array}{c} \# \text{ of F. A. accepted} \\ \text{Total \# of F. A.} \\ \text{available} \end{array}}{\begin{array}{c} \# \text{ of U. A. rejected} \\ \text{Total \# of U. A.} \\ \text{available} \end{array}} - 1$$

1-C (for  $C \geq 0$ )

Agree response set indicated a preference for accepting favorable adjectives to achieve self-image, while disagree set indicated a tendency for the rejection of unfavorable items. Actually, of course, both sets are generated in the testing situation, and the over-all agree or disagree set score for each individual indicated the relative difference in magnitude between the two and the direction of the stronger tendency.

Further, by selecting adjectives which reflected social desirability we were attempting to control for a second dimension of response set. If set is modifiable, we must account for the changes in acquiescence apart from those of the desirability continuum, which constituted the content score for this study.

Each sex group was divided at the median score of the agree set response on the first (real-self image) test. (The median cases for both groups were within the frequency of zero scores). This gave four groups:

Male—Agree Set—MAS—(N = 18)

Male—Disagree Set—MDS—(N = 17)

Female—Agree Set—FAS—(N = 47)

Female—Disagree Set—FDS—(N = 48)

F and T tests were carried out for set and content modifications within each of the female groups as a function of changes in test directions; that is, to give a real and then an ideal self. The Wilcoxon sign-rank test of

differences and the Mann-Whitney U tests were employed for men because of the smallness and skewness of the samples.

### C. RESULTS AND DISCUSSION

Both sexes gave a significantly better concept of self, as required by the changes in directions. The results for the women are given on Table 1. The FAS difference (28.65) and FDS difference (28.66) were both significant at the .01 level. The Wilcoxon values for the MAS ( $T = 0$ ) and MDS ( $T = 1$ ) were also significant at the .01 level.

Further, for both men and women, there were no significant differences in content scores between the agree and disagree set groups within the real self-concept, or the ideal self-concept. Referring to Table 1 for women, we find that the differences between the FAS and FDS groups within the "real-self" directions (.94) and "ideal-self" (1.20) were not significant. For men, Mann-Whitney U's were computed, and the results are given in Table 2. The results were similar to those obtained for the women. Thus the direction of the agree-disagree dimension of set, in and of itself, did not appear to lead to significant differences in the magnitude of the self-concept.

TABLE 1  
DIFFERENCES IN FEMALE CONTENT SCORES BETWEEN REAL AND IDEAL SELVES

Group	Mean content scores real self	S.E.	Mean content scores ideal self	S.E.	<i>t</i>
FAS	57.40	2.07	86.05	1.65	10.8*
FDS	56.46	3.60	87.25	1.15	8.2*

\*  $P < .01$

TABLE 2  
DIFFERENCES IN MALE CONTENT SCORES BETWEEN REAL AND IDEAL SELVES

Group	Real self	Ideal self
MAS	$R_1 = 268.$	$R_1 = 312.5$
MDS	$R_2 = 319.$	$R_2 = 257.5$
Mann-Whitney U	140.0(ns)	164.5(ns)

For women,  $F$  tests for set variance within each group were significant, as shown in Table 3. The  $t$  tests for differences in set magnitude, however, were *not* significant, as seen in Table 4.

The results indicated significant variation (i.e., modifiability) within the



female group. But the direction did not appear to be consistent, and this can be shown even further by the following analysis of the female group: considering only the direction of the changes, but not magnitude, of the 47 Ss in the agree set group, 22 showed a trend toward agree set, 24 were disagree set trends, and one showed no change at all; for the disagree set group, 24 showed agree set gains and 24 showed changes in a disagree set direction.

TABLE 3  
FEMALE RESPONSE SET VARIANCES BETWEEN REAL AND IDEAL SELVES

Group	Real self	Ideal self	F
FAS	(14.76) <sup>2</sup>	(49.65) <sup>2</sup>	11.3*
FDS	(16.75) <sup>2</sup>	(53.10) <sup>2</sup>	10.0*

\*  $P < .01$

TABLE 4  
DIFFERENCES IN MAGNITUDE OF FEMALE RESPONSE SET SCORES BETWEEN REAL AND IDEAL SELVES

Group	Mean response set real self	S.E.	Mean response set ideal self	S.E.	T
FAS	15.38	2.14	20.07	7.24	0.6 (ns)
FDS	-26.65	2.44	-26.65	7.74	0.0 (ns)

As seen in Table 5, for men the changes in set for both agree and disagree set response groups were in the positive direction and significant.

TABLE 5  
DIFFERENCES IN MAGNITUDE OF MALE RESPONSE SET SCORES BETWEEN REAL AND IDEAL SELVES

Group	Wilcoxon T
MAS	T = 38.5 ( $P < .05$ )
MDS	T = 14.0 ( $P < .01$ )

Paralleling our analysis of the female group above we find that for the males, of the 18 Ss in the agree set group, 14 showed trends toward agree set changes, and four were disagree set; while in the disagree set group, 14 showed trends toward agree set gains, and 3 were toward the disagree set.

The results seem to support the hypothesis and verify the findings of an earlier and limited study. Men showed a significant agree set change in response set regardless of the direction of the first set score. Women showed changes in response set, but did not follow a pattern. It would appear thus

that response set is at least modifiable within a given testing situation and differences in modifiability may exist between the sexes.

#### D. SUMMARY

The Gough Adjective Checklist was administered twice within a month to 95 women and 35 men. On the first administration they were asked to give "real self" evaluation and on the second they gave an "ideal self" image.

The hypothesis that the agree-disagree dimension of response set is modifiable and that women, unlike men, do not follow a predictable pattern of response set modification was confirmed. Men showed a tendency to assume an agree set regardless of the initial tendency, while the women generated about equal tendencies in both directions.

#### REFERENCES

1. FREDERICKSEN, N., & MESSICK, S. Response set as a measure of personality. *Educ. Psychol. Meas.*, 1959, **19**, 137-57.
2. JACKSON, D. N., & MESSICK, S. Content and style in personality assessment. *Psychol. Bull.*, 1958, **55**, 243-52.
3. SIEGAL, S. *Nonparametric Statistics*. New York: McGraw-Hill, 1956.

*Department of Psychology*  
*Utah State University*  
*Logan, Utah*



## SENSORY STIMULATION AND SECONDARY REINFORCEMENT\*<sup>1</sup>

Rutgers University

MICHAEL P. ANDRONICO AND DONALD G. FORGAYS

### A. INTRODUCTION

Until recently, studies concerned with motivational variables have emphasized largely the influence on behavior of biological drives such as hunger and thirst and the importance of primary reinforcers such as food, water, and the like, in the modification of behavior. Finding difficulty in employing these "primary" drives and reinforcers to explain at least some forms of behavior, learning theorists proposed the further explanatory principle of secondary reinforcement. According to this notion, previously neutral stimuli can acquire motivational status and become reinforcing in their own right by being associated with one or more primary reinforcers. Thus, for example, if one pairs a stimulus such as a sound or a light-flash with food in a typical learning situation, the sound or light may later be sufficient to maintain a high response rate in the absence of the food. The assumption is made, and quite often untested, that the sound or light cue itself is not able to elicit or maintain a sustained rate of the response studied.

Much research in secondary reinforcement has been concerned principally with establishing the validity of the concept itself (Bugelski, 1938; and many others). Quite recently, relevant research has examined some of the experimental conditions under which the phenomenon can be produced. Myers (11) reviewed research in the area and discussed two prominent theoretical attempts to explain secondary reinforcement, the discriminative stimulus approach (e.g., 13) and the discrimination hypothesis (1, 2). Each of these explanations assumes the validity of secondary reinforcement but the studies on which they are based have not tested satisfactorily the neutrality of the stimuli which are used as secondary cues, as Myers (11) has pointed out.

It would appear that another explanation of secondary reinforcement is

\* Received in the Editorial Office on May 11, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This report in large part is based upon a thesis study done by M.P.A. under the direction of D.G.F. and submitted in partial fulfillment of the requirements for the M.A. degree at Rutgers University. The study was supported in part by a research grant to D.G.F. from the Research Council of Rutgers University.

available, one which is considerably more parsimonious than those referred to above. It is possible that the so-called "neutral" cue, typically auditory or visual stimuli, is not neutral but may have reinforcement potential of its own. Several research studies have been reported recently which provide considerable evidence concerning the non-neutrality of sensory stimuli as reinforcing agents (8, 12, 5, 6, 7, 9, 10 and others). Butler (4) has reviewed some of these studies, most of which point to the effectiveness of sensory stimuli, such as those used in secondary reinforcement research, as reinforcers even though they have had no obvious previous association with the traditional "primary" reinforcing stimuli. Since none of the experiments in secondary reinforcement controlled sufficiently well the neutrality of the stimulus used, it seems advisable that this issue be clarified further before more elaborate theoretical models to explain the secondary reinforcement phenomenon are considered.

The present experiment is an attempt to test the reinforcement neutrality of a stimulus typically employed in secondary reinforcement studies, in this case an auditory signal. On the basis of recent evidence, our expectation is that this sensory cue will not be found to be neutral but, rather, reinforcing in its own right.

## B. METHOD

### 1. *Subjects*

The subjects were 109 albino rats from the Rutgers University Department of Psychology strain; 51 were male and 58 female. Animals were about 95 days of age at the beginning of testing and were experimentally naive at that time. They were assigned at random to the various groups comprising the study; each group had approximately an equal number of male and female animals.

Of the 109 animals which began acquisition trials, 92 (45 males and 47 females) were included in the acquisition data analyses; of the remaining 17, 4 died during these trials and 13 were eliminated because they did not respond in at least half of the acquisition trials, the criterion employed for inclusion in analysis. Of the 92 animals whose data were included in acquisition analysis, 24 were eliminated from extinction data analysis because of a planned modification in the extinction trials of the replication design of this study. This left 68 animals in the extinction analyses reported below.

### 2. *Apparatus*

The testing area consisted of an enclosure contained in an old 6.9 cubic foot refrigerator; it was 1 foot wide, 1 foot deep, and 1 foot high. The



floor and two walls of the enclosure were formed by the inside of the refrigerator, a third wall was a hinged transparent plastic sheet used as an inside door, and the ceiling was a metal grid. The fourth wall was of hardboard. Approximately in its center, two inches above the floor, extended a  $2 \times 2\frac{1}{2}$  inch food cup and, two inches to its left, a  $2\frac{1}{2} \times 4$  inch transparent lever. Depression of this bar activated a counter mounted outside of the refrigerator. Also mounted here was an automatic food-pellet dispensing mechanism which was connected by plastic tube through the back of the refrigerator to the food cup within and a Hewlett-Packard audio oscillator set at 6000 cycles per second, which in turn was connected to a  $2 \times 4$  inch speaker suspended directly above the grid ceiling of the test enclosure. The sound produced inside the test enclosure was measured by a General Radio auditory intensity meter (Model 1551-B) to be 86 to 98 decibels, varying with distance from signal source. A 40-w bulb was also suspended above the grid ceiling of the test enclosure.

Several switches were wired outside the refrigerator which, when properly set, allowed various combinations of "reinforcement" (food, sound, light) to be received in the test chamber when the lever was depressed. Appropriate animals received a single food pellet for each distinct press; light or sound, however, continued as long as the lever was depressed.

With the door of the refrigerator closed, which was the routine running condition, the test chamber within was reasonably well isolated from outside sounds including those of the mechanisms referred to above. The refrigerator interior contained enough air to obviate concern over the animal's oxygen supply for the short trial session.

### 3. Procedure

Each animal was put into the test enclosure within the closed refrigerator for 10 minutes daily for 40 days, the first 25 of which were acquisition trials and the remaining 15 were extinction trials. Trials were run on six consecutive days of each seven until completed. Each animal was food-deprived for approximately 23 hours before testing and fed for about one hour after.

There were four groups of animals in acquisition testing:

- Group A. Food and Sound: 29 animals who received the sound specified above and one 0.45 mg. food pellet for bar press.
- Group B. Food: 21 animals who received food only for bar press.
- Group C. Sound: 22 animals who received sound only for bar press.
- Group D. Control: 20 animals who received no reinforcement for bar press.

These four acquisition groups were subdivided into eight groups for the extinction trials, as follows:

Group A was broken down into three subgroups:

Group 1: 7 animals who received sound for bar press.

Group 2: 6 animals who received light for bar press.

Group 3: 8 animals who received nothing for bar press.

Group B was broken down into two subgroups:

Group 4: 7 animals who received sound for bar press.

Group 5: 7 animals who received nothing for bar press.

Group C was broken down into two subgroups:

Group 6: 7 animals who received sound for bar press.

Group 7: 8 animals who received nothing for bar press.

Group D, the control group for acquisition trials, continued to receive nothing for bar press during extinction and is referred to here as Group 8. Eighteen animals completed the extinction series in this group.

Of the eight groups above, Group 1 is that most comparable to the typical secondary reinforcement group. Group 2 was included to examine the influence of a changed sensory cue from that received in acquisition trials. Group 4 was an attempt to assess the possibility of a general sensitization effect or pseudoconditioning. Group 6 assesses the continuing influence of sensory reinforcement. The remaining four groups are control observations for the various possible effects specified for Groups 1, 2, 4, and 6.

The discrepancy between the sample sizes listed above for extinction as compared with acquisition trials has already been referred to above.

The experiment was accomplished as a replication design with approximately half of the animals of each group being run in each of two test series.

## C. RESULTS

### 1. *Acquisition*

The mean response levels over the 25 acquisition trials of the four principal subject groups of this study are given in Table 1. As can be seen there, the mean response level of the food and sound group is higher than that of the food group while the response level of the sound group is higher than that of the control group; both the food and sound and food groups responded at a level considerably higher than that of the sound and control groups. Mean response levels of these four groups of subjects for each of the 25 acquisition days are shown in Figure 1; these are the data used for acquisition analysis since the over-all trend comparisons among groups were of greatest interest here.



TABLE 1  
MEAN RESPONSE LEVELS OF THE MAIN GROUPS FOR ACQUISITION TRIALS

Group	N	Mean	Standard deviation
Food & sound	29	1015.97	409.83
Food	21	881.48	307.35
Sound	22	84.91	27.53
Control	20	65.35	27.78

These response data were not normally distributed and therefore a non-parametric technique, the sign test (14), was employed for analysis throughout. In addition, a one-tail test of significance was used for all comparisons

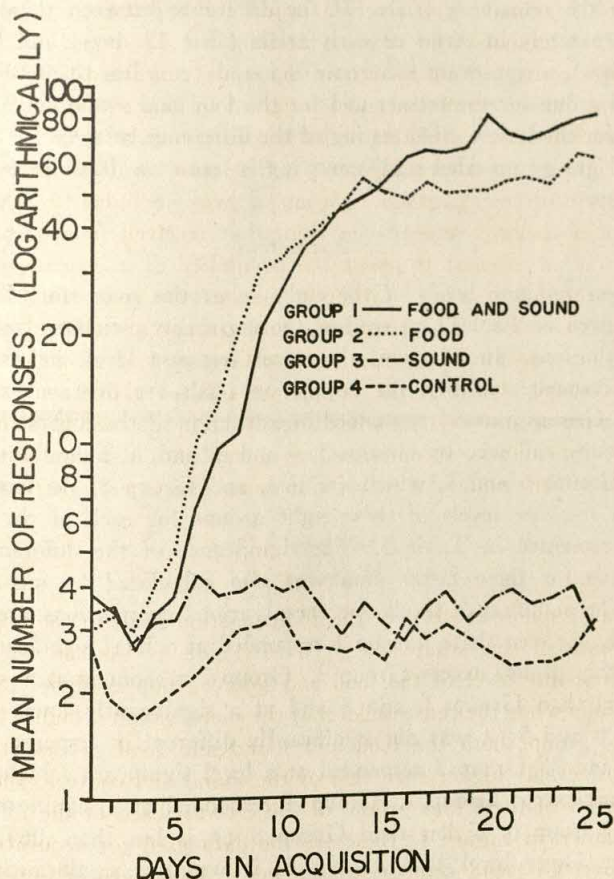


FIGURE 1  
MEAN RESPONSE LEVELS OF THE FOUR GROUPS OF SUBJECTS FOR EACH OF THE ACQUISITION TRIALS

since a general directional prediction was made that those groups receiving sensory reinforcement (sound or light) would respond at a level higher than that of groups not receiving such cues for responding.

Sign testing of the mean response levels for the 25 acquisition trials revealed a difference in favor of the food and sound and the food groups over both the sound and control groups, at the .001 level of confidence. A similar analysis comparing the food and sound group and the food group does not attain statistical significance. An inspection of Figure 1 indicates that the food group responded consistently at a level higher than that of the food and sound group over the early portion of these trials whereas the reverse is true for the remaining trials. If the difference between these groups is analyzed separately in terms of early trials (first 13 days) and later trials (last 12 days), a significant difference is found ( $p$  is less than .01) in favor of the food group for the former and for the food and sound group ( $p$  is less than .01) for the latter. Sign testing of the difference between the sound and the control groups revealed significance ( $p$  is less than .001) in favor of the sound group.

## 2. Extinction

The mean response levels of the eight subgroups over the 15 extinction trials are given in Table 2, as well as the variability estimates computed for these distributions. In addition, the mean response levels of these groups and reinforcement condition for acquisition trials are provided. As can be seen there, Groups 3 and 1 responded in extinction at the highest level of the various groups, followed by Groups 5, 4 and 2, and, at a considerably lower level, by Groups 6 and 7, which are tied, and Group 8, the lowest of all. The mean response levels of these eight groups for each of the extinction trials are presented in Table 3. The significance of the differences among these groups for these latter data was also calculated by one-tailed sign testing. The significance levels for these various comparisons are presented in Table 4. As seen there, Group 1 responded at a level significantly higher than all other groups except Group 3. Group 2 responded at a significantly higher level than Groups 6 and 8 and at a significantly lower level than Groups 1, 3 and 5; it was not significantly different in response level from Groups 4 and 7. Group 3 responded at a level significantly higher than all other groups except Group 1. Group 4 responded at a significantly higher level than Group 6, higher than Group 8 ( $p$  is less than .06), and at a significantly lower level than Group 5; it was not significantly different from Group 7. Group 5 responded at a significantly higher level than



TABLE 2  
 MEAN RESPONSE LEVELS OF THE EIGHT SUBGROUPS FOR EXTINCTION TRIALS AND SOME  
 ACQUISITION DATA FOR THE SAME GROUPS

	N	Extinction			Acquisition	
		Type of reinforcement	Mean response	Standard deviation	Type of reinforcement	Mean response
Group 1	7	Sound	195.57	26.10	Food and sound	1352.00
Group 2	6	Light	108.67	15.70	Food and sound	1047.67
Group 3	8	Nothing	197.38	37.80	Food and sound	1107.25
Group 4	7	Sound	147.71	17.83	Food	996.14
Group 5	7	Nothing	167.71	30.46	Food	1065.86
Group 6	7	Sound	55.00	26.27	Sound	71.57
Group 7	8	Nothing	55.00	17.86	Sound	82.75
Group 8	18	Control	44.47	20.37	Control	66.56

Groups 6, 7 and 8. Group 6 responded at a level significantly higher than Group 8 and significantly lower than Group 5; it was not different from Group 7. Group 7 responded at a significantly higher level than Group 8.

TABLE 3  
MEAN RESPONSE LEVELS OF THE EIGHT SUBGROUPS FOR EACH EXTINCTION TRIAL

N Group	7 1	6 2	8 3	7 4	7 5	7 6	8 7	18 8
Days								
1	48.7	27.7	44.9	34.9	40.6	4.3	3.4	3.0
2	28.9	13.2	26.8	20.9	15.9	3.9	4.1	2.9
3	18.4	10.8	22.3	14.0	17.0	3.1	2.0	3.2
4	15.6	7.0	15.5	12.0	12.1	3.3	3.4	2.6
5	12.7	6.8	9.9	9.0	10.3	5.3	3.4	2.7
6	11.3	5.2	13.0	9.7	8.9	3.7	3.6	3.6
7	9.1	3.2	11.4	6.7	7.3	4.3	4.3	3.2
8	6.6	6.0	6.8	6.1	6.4	3.9	4.0	2.4
9	8.0	6.3	7.1	5.3	8.7	3.0	3.4	2.9
10	7.0	4.3	7.1	5.9	7.3	3.6	4.0	3.2
11	7.9	3.8	6.3	1.6	4.9	3.7	3.9	2.7
12	5.9	2.7	7.6	2.3	5.1	4.7	3.5	2.4
13	4.9	2.8	5.3	2.4	4.9	2.4	3.6	3.2
14	7.0	5.8	8.3	4.0	5.4	3.3	4.0	3.2
15	4.9	3.0	5.4	1.7	5.1	2.7	4.5	3.4

TABLE 4  
SIGNIFICANCE LEVELS OF SIGN TESTS BASED UPON COMPARISONS AMONG THE EIGHT SUBGROUPS IN EXTINCTION

Groups	2	3	4	5	6	7	8
1	<.001	—	<.001	<.05	<.001	<.001	<.001
2		<.001	—	<.001	<.01	—	<.001
3			<.001	<.05	<.001	<.001	<.001
4				<.01	<.05	—	<.06
5					<.001	<.001	<.001
6						—	<.02
7							<.001

#### D. DISCUSSION

The results of the several analyses reported above support in general the notion that auditory stimulation, such as that used here, is not neutral with respect to reinforcement value. It will be recalled that the group which received only the sound cue for lever press in acquisition trials responded at a level significantly higher than that of the control group which did not receive such a cue. Moreover, for the second half of the acquisition trials the food and sound group responded at a significantly higher level than the group which received only food for lever press. Why this difference



was not found for the earlier trials, but rather the reverse relationship, is difficult to explain. It appears that such sensory stimulation may operate as disruptive stimuli when introduced in association with important primary reinforcers such as food. Continued exposure may then lead to some kind of adaptation permitting the sensory cue to operate as a positive reinforcer, which is then reflected in the higher level of lever-press response.

The extinction results lend further support to the notion that sensory stimuli are not neutral cues. Both Groups 6 and 7 which had received sound for bar-pressing during acquisition trials responded in extinction trials at a level significantly higher than Group 8, the control condition. On the basis of the comparison of these groups for both acquisition and extinction trials, it appears clear that the auditory cue alone operates as an effective reinforcer, in the operational sense, under the conditions of this study. The fact that Groups 6 and 7 do not differ in response level for extinction trials is puzzling, since Group 6 continues to receive the auditory cue for lever-press whereas Group 7 does not. It is possible that this may reflect sampling error; an inspection of Table 2 reveals that Group 7 had responded in acquisition trials at a somewhat higher level than had Group 6. Thus the extinction response levels actually reflect a greater decrease in response for Group 7 as compared with Group 6, though not a significant one. It is also possible that such responding for sensory reinforcement is highly resistant to extinction and that such differences as anticipated would not be found unless a longer series of extinction trials had been employed. Further study is required to clarify this problem.

Further information on the possibility that responses for sensory reinforcement may be highly resistant to extinction is provided by a comparison of the response levels in extinction of Groups 1 and 3 and Groups 6 and 7. Those groups which received sound for responding in acquisition trials (Groups 3 and 7) but did not receive sound for responding in extinction trials did not respond at a level different from comparable acquisition groups which continued to receive the sound stimulus (Groups 1 and 6). On the other hand, both Groups 1 and 3 respond in extinction trials at a level significantly higher than Group 5, and Groups 6 and 7 at a level significantly higher than Group 8. These comparisons indicate clearly that the influence of sound received during acquisition trials continues during extinction trials and suggest rather strongly that the effect is quite resistant to extinction.

Comparison of the response levels in extinction of Group 2 with Group 3 and Group 4 with Group 5 is of interest. Both Groups 2 and 4 received a sensory stimulus for extinction responses but not for acquisition responses.

The extinction response level of both these groups is appreciably lower than that of appropriate comparison groups (Groups 3 and 5, respectively). Since these differences were in a nonpredicted direction and a one-tailed test of significance was used throughout, it is not appropriate to discuss them as significant results. We can suggest, however, that such sensory stimuli may have a disruptive effect upon response level under certain conditions, which is in agreement with the early disruptive effect described above for the food and sound group during acquisition trials. Just what these conditions may be remains for further research to investigate.

Most studies reporting the utility of sensory cues as reinforcers employed visual stimuli, and changing stimulation in some cases. In the present experiment a constant auditory stimulus was employed. Further experimentation is indicated in examining the efficacy of other sensory cues and perhaps in changing the patterns of these stimuli in designs similar to those of the present study.

The results of this experiment, then, confirm the notion that an auditory stimulus, such as that employed here, is not neutral with respect to reinforcement value under the conditions of this study. To the extent that this cue is representative of those commonly used as secondary reinforcing stimuli in contemporary experimentation, the suggestion is made that at least some of the effect attributable to a cue on the basis of an acquired reinforcement potency gained through association with more primary reinforcers, may rather be due to a reinforcement potency which the cue already possesses.

#### E. SUMMARY

The present study examined the reinforcement neutrality of a stimulus, an auditory cue, which is commonly used in secondary reinforcement experiments.

Acquisition and extinction of a lever-press response of rats was studied by varying systematically the type of reinforcing cue received for response in both the acquisition and the extinction trials.

The results indicate clearly that so-called neutral cues such as that used here are not neutral with respect to reinforcement value. The relevance of these data as a source for clarifying some of the confusion in contemporary investigations of the phenomenon of secondary reinforcement is discussed, and the suggestion is offered that similar parametric studies should be undertaken to explore further this problem.



## REFERENCES

1. BITTERMAN, M. E., FEDDERSEN, W. E., & TYLER, D. W. Secondary Reinforcement and the Discrimination Hypothesis. *Amer. J. Psychol.*, 1953, **66**, 456-464.
2. BITTERMAN, M. E., TYLER, T., & ELAM, C. A further study of secondary reinforcement and the discrimination hypothesis. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 381-384.
3. BUGELSKI, R. Extinction with and without sub-goal reinforcement. *J. Comp. Psychol.*, 1938, **26**, 121-134.
4. BUTLER, R. A. Acquired Drives and the Curiosity-Investigative Motives. Pages 144-176 in: *Principles of Comparative Psychology*, edited by Waters, R. H., Rethlingshafer, D., & Caldwell, W. E. New York: McGraw-Hill, 1960.
5. FORGAYS, D. G., & LEVIN, H. Learning as a function of change of sensory stimulation in food-deprived and food-satiated animals. *J. Comp. & Physiol. Psychol.*, 1958, **51**, 50-54.
6. ———. Discrimination and reversal learning as a function of change of sensory stimulation. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 191-194.
7. ———. Learning as a function of change of sensory stimulation: distributed vs. massed trials. *J. Comp. & Physiol. Psychol.*, 1961, **54**, 59-62.
8. KISH, G. B. Learning when the onset of illumination is used as reinforcing stimulus. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 261-264.
9. LEVIN, H., & FORGAYS, D. G. Learning as a function of sensory stimulation of various intensities. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 195-201.
10. ———. Sensory change as immediate and delayed reinforcement for maze learning. *J. Comp. & Physiol. Psychol.*, 1960, **53**, 194-196.
11. MYERS, J. L. Secondary Reinforcement: A review of recent experimentation. *Psychol. Bull.*, 1958, **55**, 284-301.
12. PREMACK, D., COLLIER, C., & ROBERTS, C. L. Frequency of light-contingent bar pressing as a function of the amount of deprivation for light. *Amer. Psychol.*, 1957, **12**, 411 (Abstract).
13. SHOENFELD, W., ANTONITIS, P. J., & BERSH, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. Exp. Psychol.*, 1950, **40**, 40-45.
14. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Department of Psychology  
Rutgers University  
New Brunswick, New Jersey

## CONTOUR ABSENCE AS A CRITICAL FACTOR IN THE INHIBITION OF THE DECAY OF A MOVEMENT AFTEREFFECT\*<sup>1</sup>

*Western Psychiatric Institute, University of Pittsburgh, School of Medicine*

---

IRWIN M. SPIGEL

---

### A. INTRODUCTION AND PROBLEM

Previous research (2) had indicated that the interpolation of an interval of darkness between *S*'s exposure to a rotating spiral and re-exposure to the stationary pattern had brought about the inhibition of the normal decay of the negative aftereffect of such movement stimulation. This inhibitory manifestation was viewed as of potential value as a perceptual probe in the psychophysiological investigation of perseverative neural processes. In addition, the author (3) has employed the inhibitory effect under interocular procedures to demonstrate support for a position of central mediation in aftereffect phenomena—a critical question with regard to the more general problem of ascertaining the nature of neural events underlying aftereffect perception.

It appeared that the experimental manifestation of aftereffect inhibition constituted a perceptual analogue of Thompson and Bryant's (6) and Thompson's (5) demonstrations of perseverative activity after new learning when darkness was interpolated between training and test.

Specifically, the current experiment was designed to determine whether the sufficient condition for the inhibition of the aftereffect decay was the absence of illumination in the post-movement interval, or the absence of interpolated contour following exposure to rotation.

### B. METHOD

#### 1. *Apparatus*

A detailed description of the equipment and wiring is included in unpublished portions of the author's doctoral dissertation (4). "A guillotine" assembly was constructed so as to allow a white wood screen to fall in place

---

\* Received in the Editorial Office on May 14, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Based on portions of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Temple University, January 1962. The author wishes to thank the chairman, Dr. Hughbert C. Hamilton, and members of his committee, Drs. William J. Carr and Horace A. Page, with all of whom it has always been a pleasure to work.



in front of the primary target when a set of solenoid shafts were withdrawn. The screen could fall a second distance, again revealing the primary target when a second set of solenoids were activated. The primary target itself consisted of a white 4-inch disc on which were drawn six black curved contours radiating from the center. The disc was fitted to the shaft of a 10 rpm motor. Only the disc and "guillotine" assembly were visible to *Ss* during the experimental trials. These were placed on a table of standard height in a light-proof room. Gray curtains and screens were placed so as to keep *E* and the operations console containing switches, timers and chronoscope from *S's* view. *S* sat at a viewing distance of 32  $\frac{3}{4}$ " from the target, so that the disc presented a visual angle of 15° 14' at the eye. Distance from *S* to disc was kept constant by way of a fixed chin rest. *S* was also in control of a toggle switch which could stop the chronoscope when flipped upward. Motor, timers and chronoscope were wired so that five different sequences of events were possible.

## 2. Subjects

Twenty-five male students in introductory psychology courses were selected and assigned so as to yield five separate and equated groups with reference to the amount of aftereffect perceived after 30 seconds of disc rotation. There were five *Ss* in each group.

## 3. Procedure

A multiple Latin square design, described by Edwards (11, pp. 319-326) was employed. There were five separate  $5 \times 5$  squares of the same treatment pattern.

Initially, each of the 25 *Ss* selected were given five 30-second exposures to the rotating design, and the duration of his aftereffect recorded. Brightness of 2.03 millilamberts was provided by the combined illumination of 75-watt and 15-watt lamps. *S* had been instructed to fixate on the center of the disc during rotation, and to flip his switch upward when he no longer perceived the aftereffect. This operation stopped the chronoscope which had started at the end of disc rotation. Since it was necessary that each *S* employed perceive the negative aftereffect, a special trial preceded this series in which *S* was instructed merely to describe any movement he saw following the termination of disc rotation. Actual instructions read to *Ss* may be found in Spigel (4). All *Ss* reported the aftereffect. In the actual series, *S* was given six trials, of which only the last five were employed in computing a mean aftereffect duration. It was on the basis of these means that assignment

to groups was made. Experimental trials began one week later. The five conditions comprising the square were as follows:

*Condition A:* This was essentially a replication of the determination of *S*'s mean aftereffect, as in the preliminary series.

*Condition B:* Residual aftereffect duration was obtained after an interval of darkness equal to the *S*'s preliminary mean aftereffect duration was interpolated following the cessation of rotation.

*Condition C:* An interval of relatively unpatterned light, as reflected from the white screen of the guillotine assembly, which had fallen into place in front of the disc at the termination of rotation, was interpolated and the residual aftereffect duration recorded.

*Condition D:* This was essentially the same as Condition *C*, except that instead of the brightness remaining equal to that of the disc itself, the illumination on the screen during the interpolated interval was reduced to .14 millilamberts. As before, the residual aftereffect was recorded.

*Condition E:* This situation imposed an interrupted sequence of events on *S*. Following the 30-second rotation of the disc, *S* was allowed to view the reverse movement of the stationary disc for an interval equal to one-half of his mean preliminary aftereffect. *S* was then placed in darkness, at the termination of which, illumination was restored and residual aftereffect recorded.

The five conditions were presented to *Ss* on successive days. Five trials were administered for each condition. The duration of each *S*'s residual aftereffect following each of the interpolated conditions was read from the chronoscope after the individual trials and a mean computed from the last four.

### C. RESULTS

The mean aftereffect durations for *Ss* assigned to the five Latin squares, as based on the preliminary series of trials, were 9.4, 9.2, 9.4, 8.8 and 10.3, respectively. The failure of the *F* ratio ( $F = .08$ ,  $df = 4/20$ ) between *Ss* in each of the five squares to attain statistical significance was taken to indicate equation of the squares as regards amount of aftereffect perceived by assigned *Ss*.

Table 1 contains an analysis of the variance in the five Latin squares for the duration of residual aftereffect reported by *Ss* following differential post-exposure stimulation. The *F* ratio of 14.88 ( $df = 4/80$ ) for aftereffect duration was significant at less than the .01 level. *F* ratios for order and sessions were not statistically significant.



TABLE 1  
ANALYSIS OF VARIANCE IN THE FIVE LATIN SQUARES FOR THE ACTUAL DURATION  
OF THE MOVEMENT AFTEREFFECT FOLLOWING DIFFERENTIAL  
INTERPOLATED STIMULATION

Source of variation	Sum of squares	df	Mean square	F	P
Independent observations:					
Order of presentation	337.76	4	84.44	1.36	> .05
Residual between individuals (error)	1215.76	20	60.78		
Total between individuals	1553.52	24			
Correlated observations:					
Differential interpolated stimulation	176.21	4	44.05	14.88	< .01
Experimental sessions	9.64	4	2.41	.81	—
Residual from Latin square (error)	65.74	12	5.48	1.85	> .05
Residual from within individuals (error)	237.59	80	2.96		
Total within individuals	489.18	100			
Total for experiment	2042.70	124			

Significance levels of differences between the means for residual aftereffect duration under the five conditions of the experiment are contained in Table 2. The differences between the mean aftereffect determined for Condition *A* and the four other conditions are statistically significant at less than the .01 level in all cases. Mean difference in aftereffect under Conditions *C* and *D* (interpolated unpatterned illumination at different intensities) was also statistically significant at less than the .05 level. Differences between the mean aftereffect under the other experimental conditions were not statistically significant.

#### D. DISCUSSION

The experiment was concerned with the question of the sufficient condition for the inhibition of the decay of a movement aftereffect following visual stimulation by repetitive, unidirectional objective pattern movement.

If none of the interpolated stimulus conditions had an inhibitory effect on the normal decay process—and since Conditions *B*, *C* and *D* included interpolated intervals equal to *S*'s mean aftereffect—then the residual aftereffect should have been zero, or close to zero, when *S*s were re-exposed to the stationary target pattern.

The broad indication from this data, however, was that all of the inter-

TABLE 2  
MEANS, DIFFERENCES BETWEEN MEANS, AND SIGNIFICANCE LEVELS OF DIFFERENCES  
FOR THE RESIDUAL AFTEREFFECT DURATION FOR *SS* IN THE LATIN SQUARES  
FOLLOWING DIFFERENTIAL INTERPOLATED STIMULATION

Interpolated condition***	Mean residual AE (in Secs.)	Mean differences between conditions			
		B	C	D	E
A. Simple aftereffect: no interpolated stimulation	9.5	2.4**	3.4**	2.1**	3.1**
B. Total darkness following objective movement	7.1		1.0*	.3	.7
C. Unpatterned illumination following objective movement	6.1			1.3**	.3
D. Unpatterned illumination of reduced intensity	7.4				1.0*
E. Exposure to contour: then darkness, after objective movement	6.4				

\*  $p < .05$

\*\*  $p < .01$

\*\*\* Interpolated intervals were equal to each *S*'s mean aftereffect duration as determined in a preliminary series of trials.

polated stimulation had brought about some inhibition of the decay process. Further, it appeared that the interpolated stimulation employed differentially affected the rate of decay of the movement aftereffect. Moreover, it seemed that visual pattern or contour deprivation represented the sufficient condition for the inhibition of the aftereffect decay.

The interpolated conditions included darkness, relatively unpatterned illumination of two different intensities, and successive exposure to a stationary pattern and an interval of darkness. What these interpolated conditions appear to have in common is the relative absence, for a period of time, of spatially or temporally patterned visual stimulation, i.e., the relative absence of contours and contour processes from the involved portion of the visual field.

One can see from Table 2 that the duration of residual aftereffect in all conditions of interpolated stimulation was less than that which was recorded



when measurements followed immediately after the 30-second exposure to objective movement. This finding argues for some unimpeded, and perhaps constant, rate of aftereffect loss which is independent of post-exposure stimulus events. This suggests some functional dissipation or "leakage" in the neural events underlying aftereffect experience.

The fact that the residual aftereffect duration following the interpolation of unpatterned illumination of higher intensity (Condition *C*) was significantly less than that following darkness, as well as the duration following interpolated unpatterned illumination of considerably reduced intensity presents subtle support for the importance of contour processes, or their absence, in the inhibition of the aftereffect decay. That the residual aftereffect duration under Condition *C* was not statistically different from the residual aftereffect under Condition *E* (stationary pattern exposure followed by darkness) further argues for this position; as does the failure to find differences between the residual aftereffect in Condition *B* and that obtained for Conditions *D* and *E*.

The explanation may be in the apparatus itself. While the screen, painted white, which dropped into position in front of the disc in Conditions *C* and *D* was to have presented a period of "unpatterned" illumination to *Ss*, it was in fact only relatively so. The flat white applied to the screen in several coats by brush left a fine, but noticeable texture in relief on its surface. Under the conditions of illumination—high angle from behind *S*—the texture could have been sufficiently marked by the effects of shadow to have rendered the screen surface one of minimal, rather than zero, contour. That is, sufficient contour could have been presented under the illumination of the given intensity to bring about significantly more decay than under a condition of zero contour (*ganzfeld*). Or stated negatively, the minimal contour visible to *S* was not significant to bring about maximal inhibition or delay of the aftereffect decay—or those neural processes underlying the illusory experience.

Under Condition *D*, however—where the screen was interpolated under markedly reduced illumination—the amount of light flux on the target surface very likely was inadequate to produce sufficient texture to give rise to contour processes of measurable influence. This question will be explored in future work.

The possibility that the effects of contour, as facilitating decay of the apparent movement experience, act similarly in time and space is also suggested. The fact that the residual aftereffect in Condition *E* is significantly less than that under Condition *D*, but not statistically different from the aftereffect duration in Condition *C* would also support such interpretation.

That is, the relative absence of visible contour under the reduced illumination of the screen in Condition *D* could have, as stated earlier, made for near maximal decay inhibition. Under Condition *E*, *Ss* were exposed to the stationary contour in full illumination for a period of time equal to one-half the duration of their aftereffect before being placed in darkness—ample time for neural events underlying the illusory movement to undergo some degree of dissipation. An explanation of the decay under Condition *C* has already been offered. Now, it remains a distinct possibility that the rate of decay of the movement aftereffect is a function of some additive or multiplicative interaction of events in time and space. It could be that under the circumstances of the experiment, the processes underlying the decay of the aftereffect for the shorter period of time with the more prominent contour (the stationary pattern itself) as in Condition *E*; and for the longer period of time with the less marked contour (the textured surface of the screen) in Condition *C*, were roughly equivalent—yielding no statistical difference in the residual aftereffect duration.

The determination that residual aftereffects following interpolated darkness (Condition *B*) were not statistically different from the residual duration of apparent movement in Condition *D* may be explained in the same terms. In both cases the absence of contour remains a common stimulus dimension.

That the residual aftereffects are not statistically different in Conditions *B* and *E* further emphasizes complexities in the rate of decay of the events underlying the illusion. It appears that in a consideration of interpolated stimulation, the *change* of stimulation itself is a critical factor in the processes forming the basis of the aftereffect experience and of the inhibition or delay of this phenomenal movement. That is, the sudden change in objective conditions may itself introduce collative processes that augment arousal input into the central systems underlying the phenomenon. In Condition *E*, for example, the inclusion of a period of exposure to the stationary pattern prior to the interval of darkness necessitated one additional abrupt change in stimulus events over those in the other conditions of the experiment.

The two broad inferences that emerge, then, may be restated as follows: (*a*) the normal rate of decay of the aftereffect of visually perceived movement may be inhibited by the interpolation of conditions which markedly reduce or eliminate contour in the involved portions of the visual field; and (*b*) that the *change* in stimulus condition may itself be an active factor in the mediation of the aftereffect and the inhibitory process presumed to function under the stated conditions.



## E. SUMMARY

An experiment was designed to determine the sufficient condition for the inhibition of the normal rate of decay of a movement aftereffect. In addition to a regular aftereffect determination, conditions of darkness, relatively homogeneous illumination of two different intensities, and the successive exposure to a stationary pattern and darkness were interpolated between the termination of objective movement and re-exposure to the stationary pattern. A determination of *S*'s normal aftereffect was also obtained. Twenty-five male *S*s were employed in a  $5 \times 5$  multiple Latin square design. Mean residual aftereffects were computed for all *S*s under each condition. All conditions of interpolated post-movement stimulation succeeded in bringing about inhibition of the decay of the aftereffect, but in varying degrees. Maximal inhibition was evident under the conditions of interpolated darkness and illumination of reduced intensity. The two conclusions from the experiment were: (a) the normal rate of decay of a movement aftereffect may be inhibited by the interpolation of conditions which markedly reduce or eliminate contour in the involved portions of the visual field; and (b) that *change* in stimulus conditions itself is an active factor in the mediation of the aftereffect and the inhibitory process.

## REFERENCES

1. EDWARDS, A. E. *Experimental Design in Psychological Research*. New York: Holt, 1950.
2. SPIGEL, I. M. The effects of differential post-exposure illumination on the decay of a movement aftereffect. *J. of Psychol.*, 1960, **50**, 209-210.
3. ———. The use of decay inhibition in an examination of central mediation in movement aftereffect phenomena. *J. Gen. Psychol.*, in press. (a).
4. ———. The effects of differential post-exposure stimulation on the decay of the aftereffect of visually perceived movement. Unpublished doctoral dissertation. Temple Univ., 1962. (b).
5. THOMPSON, R. Retroactive effect of interpolated visual stimulation. *Psychol. Rep.*, 1957, **3**, 183-188.
6. THOMPSON, R., & BRYANT, J. H. Memory as affected by activity of the relevant receptor. *Psychol. Rep.*, 1955, **1**, 393-400.

Western Psychiatric Institute  
University of Pittsburgh  
School of Medicine  
3811 O'Hara Street  
Pittsburgh 13, Pennsylvania

## PROFILES OF AN ADOLESCENT\*

*Department of Education, Oklahoma State University*

---

SAMUEL GOLDMAN<sup>1</sup>

---

### A. INTRODUCTION

Students of adolescent behavior seem to agree that the central problem for the adolescent is that of establishing a sense of identity, adolescence being that period in life during which a young person strives for clarification as to who he is and what his role in society might be. Erikson expands on this notion slightly when he suggests that not only is the adolescent seeking self-identity and role clarification, but he is also vitally concerned for how others regard him and how the assessments of others compare with his own conception. He says:

What the regressing and growing, rebelling and maturing youths are now primarily concerned with is who and what they are in the eyes of a wider circle of significant people as compared with what they themselves have come to feel they are . . . (2, p. 266)

A review of the literature on adolescence reveals a paucity of information regarding the perceptions people have of our youth. How do adults picture today's adolescent and what kind of person do they think he *should* be? How does this adult picture compare with the adolescent's picture of himself?

These questions and others formed the focus of a study recently completed at the University of Chicago (5). An examination was made of the perceptions of and disagreement among educators, non-educators, and high school students with regard to their real and ideal images of the adolescent.

### B. METHOD

The adolescent referred to in this study was the high school graduate. It was felt that respondents could better complete the instrument if the adolescent was examined at a given point in time . . . at graduation from high school.

#### 1. Sample

The sample was obtained from three school districts, with the respondents ranging from upper-middle to lower-middle in socio-economic background.

---

\* Received in the Editorial Office on May 15, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author is indebted to Dr. R. F. Campbell, Dr. P. W. Jackson, and Dr. C. Briner for their guidance during the conduct of this study.



A total of 1,554 instruments were distributed, of which 1,333 were completed and returned. Included among the respondents who returned questionnaires were 153 high school teachers, 224 non-educators (drawn from two PTA's, a High School Women's Music Booster Club, a Labor Council, and members of three men's service organizations), and 956 high school students from grades nine through twelve.

## 2. Procedure

An instrument adapted from one developed by Battle (1), was designed to include items which described the qualities normally identified with the high school graduate. One statement was used to describe each quality and the respondents were instructed to: (a) rank the statements from one to ten to describe the qualities which you feel are characteristic of today's high school graduate; and (b) rank the same statements again from one to ten to describe the qualities which you feel *should* be characteristic of an ideal high school graduate.

The ten statements were:

..... is basically spiritual in attitude toward life	(religious) <sup>2</sup>
..... feels that the prime goal in life is to make as much money as possible	(economic)
..... knows the political issues of the day and is willing to be involved in political activity	(political)
..... enjoys cultural activities and has a taste for the finer things of life	(aesthetic)
..... will sacrifice personal comfort for the comfort of others	(altruistic)
..... is able to make new friends easily	(social)
..... feels that the prime goal in life is to have as much fun as possible	(hedonistic)
..... takes great interest in developing a good figure or a strong body build	(physical)
..... has outstanding character traits such as honesty and trustworthiness	(ethical)
..... has some knowledge of many things and a desire to learn more	(theoretic)

<sup>2</sup> These categories were set up for ease of discussion. The respondents did not see these category headings and reacted only to the statements. For this reason the investigator, while using the headings in the discussion, cautions that they take their meaning from the substance expressed in each statement.

## C. FINDINGS

1. *Real and Ideal Images*

Table 1 presents the means and ranks of educators,\* non-educators and high school students on both the real and ideal image of the high school graduate.

a. *Educators.* On the real image, educators ranked the economic, social, and hedonistic as the three qualities which are most characteristic of today's graduate. The qualities which they felt were least characteristic of today's graduate were the political, aesthetic, and altruistic. On the ideal image the educators ranked the ethical, religious, and theoretic as the three qualities which should ideally characterize the graduate while the physical, hedonistic, and economic were ranked as the qualities which should ideally least characterize the graduate. The qualities which showed the widest variation in rankings between the real and ideal were the economic and the hedonistic. The hedonistic and economic qualities were third and first respectively on the real image but ninth and tenth respectively on the ideal image. Educators apparently felt that the desire for money and fun, the qualities which for them are most characteristic of today's graduate, should ideally be least characteristic of him.

b. *Non-educators.* The qualities which non-educators considered to be most characteristic of today's graduate are the social, theoretic, and the economic, while the qualities they felt were least characteristic of him were the religious, political, and the altruistic. On the ideal image non-educators ranked the ethical, religious, and theoretic as the qualities which ideally should be characteristic of the graduate while they placed the physical, economic, and the hedonistic qualities among those which should be least characteristic of him. Non-educators included the theoretic among the qualities which are characteristic of both the real and the ideal graduate. This quality was ranked second on the real and third on the ideal.

The qualities which showed widest variation between real and ideal are the religious and the economic. The religious quality was ranked eighth on the real and second on the ideal while the economic quality was ranked third on the real and ninth on the ideal.

c. *Students.* Students felt that the qualities which were most characteristic of today's graduate were the theoretic, social, and the ethical. The qualities which they felt were least characteristic of today's graduate were the political, altruistic, and physical. The qualities which were ranked among the most characteristic of the ideal graduate were similar to those on the real except



TABLE 1  
SUMMARY OF MEANS AND RANKS OF REAL AND IDEAL IMAGES<sup>a</sup> BY COMPOSITE EDUCATOR, NON-EDUCATOR AND STUDENT SAMPLE

Qualities	Real						Ideal					
	Educator (N = 134)		Non-educator (N = 203)		Student (N = 925)		Educator (N = 135)		Non-educator (N = 203)		Student (N = 926)	
	$\bar{x}$	Rank	$\bar{x}$	Rank	$\bar{x}$	Rank	$\bar{x}$	Rank	$\bar{x}$	Rank	$\bar{x}$	Rank
Religious	6.74	7	6.09	8	5.70	6	2.81	2	2.77	2	3.86	3
Economic	2.70	1	4.26	3	5.16	4	9.31	10	8.74	9	8.11	9
Political	7.26	8	7.22	9	6.37	8	5.18	5	5.70	7	5.39	6
Aesthetic	7.43	9	5.66	6	5.37	5	4.47	4	4.81	4	5.58	7
Altruistic	7.87	10	7.70	10	6.40	9	5.26	6	5.60	6	4.59	4
Social	3.28	2	3.87	1	4.26	2	5.92	7	5.18	5	4.97	5
Hedonistic	4.19	3	5.74	7	5.74	7	9.09	9	9.10	10	8.18	10
Physical	5.08	5	5.50	5	6.78	10	7.21	8	7.03	8	7.84	8
Ethical	5.70	6	4.90	4	5.06	3	2.59	1	2.70	1	3.33	2
Theoretic	4.75	4	4.06	2	4.16	1	3.16	3	3.36	3	3.14	1

<sup>a</sup> Rho between Real and Ideal for educators is —.53

Rho between Real and Ideal for non-educators is .15

Rho between Real and Ideal for students is .46

that the religious quality replaced the social quality on the ideal image. The qualities which students felt were least characteristic of the ideal graduate were the physical, hedonistic, and the economic.

The qualities which showed greatest variation in rankings between real and ideal were the economic and the altruistic. The economic was ranked fourth on the real image and ninth on the ideal. The altruistic was ranked ninth on the real image but fourth on the ideal.

## 2. Within-Group Agreement

From the data presented in Table 2, it may be seen that educators, non-educators, and high-school students show: (a) very little agreement on the real image; and (b) comparatively high agreement on the ideal image.

It seems reasonable to conclude from these data that members of these three groups have some common expectations regarding the qualities the ideal graduate *should* possess but relatively few common perceptions of the qualities today's graduate *does* possess.

For both the real and ideal images educators show the highest degree of within-group agreement while the lowest degree of within-group agreement is shown by the students. For the real image in particular educators show considerably more agreement among themselves than do either the non-educators or the students. This may be due to the professional training of the high school educators in which they are taught to identify certain qualities common to all adolescents.

Although the coefficients in Table 2 are statistically significant, the fact that they indicate such low agreement on the real image places a limitation upon the generalizations which may be made concerning the perceptions of people regarding the qualities possessed by today's graduate. Even the degree

TABLE 2  
COEFFICIENT OF CONCORDANCE: DEGREE OF WITHIN-GROUP AGREEMENT FOR THE REAL AND IDEAL IMAGES OF EDUCATORS, NON-EDUCATORS AND STUDENTS<sup>a</sup>

Group	N	Real	N	Ideal
Educators	134	.35	135	.64
Non-educators	203	.18	203	.56
Students	925	.08	926	.40

<sup>a</sup> An *F* test for significance revealed that all of the coefficients in this table are significant at better than the one per cent level.

of agreement on the ideal image, though considerably higher than that on the real image does not give strong evidence that groups agree among themselves on what qualities the ideal graduate *should* possess. This is particularly



true for high school students whose coefficient of concordance of .40 on the ideal image is still low.

One reason for low within-group agreement may be that the respondents were thinking of a specific adolescent. Students may have been thinking of themselves and adults may have been thinking of their child or of other adolescents with whom they were acquainted. Since there is great variability among adolescents one might suspect that the use of a real model on which to base the real image would lead to disagreement on the real image.

Another explanation for this low within-group agreement on the real image is suggested by the literature on adolescence. The adolescent occupies a peculiar position in American culture. Perceived as neither adult nor child he is often referred to as the "marginal man" (9, p. 41) because he holds a rather nebulous social position somewhere between childhood and adulthood. This leads to a situation in which adults are not always certain as to "what the adolescent is all about." Lacking any clear picture of him, adults often are left with a rather diffuse image of the adolescent, one which often is in conflict with the image which adolescents have of themselves.

This explanation suggests that *in fact* there is no uniformity of perception among adults and adolescents; hence a low degree of within-group agreement on a real image can be expected.

The higher within-group agreement shown for the ideal image might be traced to some agreed upon set of values which pervades American society. Although the items on the instrument are personal qualities, they also are expressions of values. The ideal image is really an expression of the values of the respondents and the values they wish the graduate to have. Thus the similarity in expectations for the ideal graduate may be due to the agreement among respondents in certain of the secular or sacred values of American society (3).

### 3. *Agreement Among Groups*

From Table 3 it may be seen that, except for the low agreement between educators and students on the real image, there is a surprisingly high degree of agreement between non-educators and students and between non-educators and educators on both the real and ideal images and between educators and students on the ideal image. These findings are surprising in light of some of the findings reported in the research regarding agreement among adults and adolescents in their perceptions. Studies by Hess and Goldblatt (6), Remmers and Weltman (10), Jenkins and Lippitt (8), and by Gilbert (4) have provided evidence to suggest that the attitudes and perceptions of ado-

TABLE 3  
DEGREE OF AGREEMENT AMONG EDUCATORS, NON-EDUCATORS AND STUDENTS  
IN THEIR RANKING OF THE REAL AND IDEAL IMAGE

Group	Real	Ideal
Educators and non-educators	.76	.94
Educators and students	.48 <sup>a</sup>	.84
Students and non-educators	.78	.88

<sup>a</sup>  $p = .1$  on a one-tailed test of significance. All other figures in this table have a  $p$  of a .01 on a one-tailed test of significance.

lescents are quite different from those of their parents and their teachers. The findings of this study indicate that adolescents and their parents agree rather highly in their perceptions of both the real and ideal image of the graduate. While adolescents and their teachers show rather low agreement on the real image they do agree quite highly on the ideal image.

With few exceptions, no systematic relationship could be found between age, sex, and level of education and the degree of within-group agreement as well as the degree of agreement among groups.

#### 4. *Level of Dissatisfaction*

Hess and Goldblatt suggest that "it is generally assumed that the attitudes of the society toward its teen-age members are characteristically depreciatory and often hostile." Based on interviews with adolescents, they make the following statements:

Our preliminary interviews with adolescents revealed their awareness of a presumed inferior reputation among adults. Adolescents frequently expressed the belief that they are, as a group, subject to condemnation, criticism, and general devaluation by adults, and that there exists among adults a stereotype of adolescents as sloppy, irresponsible, unreliable, inclined toward destructive and anti-social behavior (7, pp. 459-460).

If we accept this statement as being a fair assessment of adult perceptions we may conclude that these negative feelings are indicative of some dissatisfaction adults have for adolescents. It is probably safe to assume that adults are dissatisfied because there is a discrepancy between how they see the adolescent behaving and how they believe he *should* behave. In other words, dissatisfaction may be the result of the differences adults perceive between the real and ideal image of the adolescent.

The rank order correlations between the real and ideal image for each group are shown in Table 1. For educators, a comparison between the rankings on the real image with those on the ideal image, revealed a negative rank



order correlation of  $-.53$ . This suggests that the educators in this sample expressed extremely high dissatisfaction with today's graduate. The rank order correlation between real and ideal for non-educators was  $.15$ . While this is quite low, by contrast with educators who show high dissatisfaction non-educators show some, though slight satisfaction with today's graduate. When the two image profiles depicted by the students were compared a rank order correlation of  $.46$  was found, suggesting that of the three groups, students showed least dissatisfaction with today's graduate.

The data were further analyzed to find answers to two questions: (a) was there any between-group agreement on the level of dissatisfaction for particular qualities?; and (b) which group showed highest dissatisfaction with each quality?

The highest mean differences for all three groups were found for the economic and hedonistic qualities, thus indicating that educators, non-educators and students expressed highest dissatisfaction with these two qualities in today's graduate. Among the lowest mean differences found for all three groups were the theoretic, political, and physical qualities, indicating that of the ten qualities, educators, non-educators, and high school students expressed least dissatisfaction with these three. The political and theoretic qualities were the only two where no significant differences could be found among the mean differences for all three groups. On all other qualities significant differences for at least one pair of groups could be found.

In all cases the mean differences were higher for educators than they were for non-educators, indicating that on all qualities educators showed higher dissatisfaction with today's graduate than did non-educators. This was particularly true for the economic quality where the difference between the mean differences was  $1.85$  and the hedonistic quality where the difference in mean differences was  $1.42$ . Educators and non-educators showed almost the same degree of dissatisfaction for the theoretic, ethical, political, and religious qualities. The differences in mean differences for all the qualities but these four were significant at  $.05$  W.S.D. (11).

As was the case with the non-educators, educator mean differences were higher than those of the students on all qualities, indicating that educators were more dissatisfied with today's graduate on all qualities than were the students. The differences were particularly striking for the economic and hedonistic qualities. The differences between educators and students were lowest for the political, altruistic, and theoretic qualities. Except for these three qualities, all differences between the mean differences of educators and students were significant at  $.05$  level.

Where the mean differences of non-educators and students were compared, the differences in level of dissatisfaction between these two groups were not so great as those shown when educators were compared with either of them. The highest differences were shown for the economic quality. It was only for the religious, economic, and social qualities that the differences in mean differences between the real and ideal image were significant at .05 W.S.D.

#### D. SUMMARY

The findings may be summarized as follows:

1. Educators viewed today's graduate as being characterized primarily by an interest in making as much money and having as much fun as possible, and being able to make new friends easily. Knowledge of political issues, and a willingness to participate in political activity, a taste for cultural activities, and a willingness to sacrifice personal comfort for the comfort of others were perceived by educators as being least characteristic of today's graduate.

2. Non-educators perceived today's graduate as being primarily characterized by his ability to make new friends easily, his knowledge of many things and desire to learn more, and by his interest in making as much money as possible. Non-educators believed that today's graduate is least characterized by his willingness to sacrifice personal comfort for the comfort of others, by his knowledge of political issues and willingness to participate in political activity, and by his being basically spiritual in attitude toward life.

3. For the high school student, today's graduate is primarily interested in knowing many things, having a desire to learn more, making new friends and in being honest and trustworthy. Students perceived the graduate as being least interested in the political issues of the day and being involved in political activity, sacrificing personal comfort for the comfort of others, and developing a good figure or a strong body build.

4. Educators, non-educators, and students agreed on their rankings of the qualities which should be most and least characteristic of the ideal graduate. They thought that ideally the high school graduate should be characterized by outstanding character traits such as honesty and trustworthiness, a spiritual attitude toward life, and a desire to learn more. They believed that the graduate should be least interested in developing a good figure or strong body build, having as much fun and making as much money as possible.

5. Educators, non-educators and students showed higher within-group agreement on the ideal image than they did on the real image. For both the real and ideal images educators showed the highest degree of within-group



agreement while the lowest degree of within-group agreement was shown by high school students.

6. Educators and non-educators, and students and non-educators showed high agreement on their rankings on both the real and the ideal images. Educators and students showed higher agreement on the ideal image than they did on the real image.

7. When the educator rankings on the real image were compared with their rankings on the ideal image, a negative rank order correlation of  $-.53$  was found. The qualities which showed the widest variation in educator rankings between real and ideal were the economic and the hedonistic, the former being ranked first on the real but tenth on the ideal, and the latter being ranked third on the real and ninth on the ideal. The quality which for educators showed least variation between real and ideal was the theoretic, being ranked fourth on the real and third on the ideal.

8. A rank order correlation of  $.15$  was found between the real and ideal image profiles as perceived by non-educators. The qualities which for non-educators showed the greatest difference in ranking between real and ideal were the religious and the economic, the former being ranked eighth on the real and second on the ideal, with the latter being ranked third on the real and ninth on the ideal. The quality which showed the least difference between real and ideal was the theoretic being ranked second on the real and third on the ideal.

9. Students showed a rank order correlation of  $.46$  between their real and ideal image profiles. The qualities which showed the greatest variation between real and ideal were the economic and the altruistic, the former being ranked fourth on the real and ninth on the ideal, and the latter being ranked ninth on the real and fourth on the ideal. The theoretic quality was ranked first by students on both the real and ideal and the ethical quality was ranked third on the real and second on the ideal.

10. The mean differences between the real and ideal rankings for each quality were taken as an indication of the level of dissatisfaction shown for the graduate by educators, non-educators, and students.

a. The mean differences for all qualities were highest for educators.

b. The highest mean differences for all three groups were found for the economic and hedonistic qualities, indicating that educators, non-educators, and students expressed highest dissatisfaction with these qualities in today's graduate. Among the lowest mean differences found for all three groups were the theoretic, political, and physical qualities indicating that of the ten qualities, educators, non-educators, and students expressed least dissatisfaction for these

three. The political and theoretic qualities were the only two where no differences which were significant at the .05 level could be found among all three groups.

c. For all ten qualities the mean differences were higher for educators than they were for non-educators, indicating that on all qualities educators showed higher dissatisfaction with today's graduate than did non-educators. The differences between educator and non-educator mean differences for all qualities but the theoretic, ethical, political, and religious were significant at the .05 level.

d. The differences between educators and students on all their mean differences were lowest for the political, altruistic, and theoretic qualities. Except for these three qualities, all differences between the mean differences of educators and students were significant at the .05 level.

e. A comparison between the mean differences of non-educators and students revealed that on all qualities but the altruistic and the physical, non-educator mean differences were higher than those of the students. The differences in their mean differences on the religious, economic, and social qualities were the only ones significant at the .05 level.

11. There was little systematic relationship between the rankings on both the real and ideal images and the variations in age, sex, or level of education.

The data suggest that adults (educators and non-educators) appeared to be more dissatisfied with today's graduate than were the adolescents. From the adolescent point of view, the qualities which he exhibited are not too unrelated to what he should ideally exhibit. Adults, particularly educators, do not perceive today's graduate in such a favorable light, suggesting that he has, in their opinion, a long way to go before he can approach what they consider to be ideal.

This difference in level of dissatisfaction emphasizes the conflict which exists between these two groups. What lies behind this conflict may, to some extent be inferred from their differences in the rankings on the religious, ethical and theoretic qualities on the ideal image.

The ethical and religious qualities deal with such traits as honesty, and trustworthiness, and a spiritual attitude toward life. They are part of the inner conscience which serves as the control valve to keep the individual from going against the accepted moral and ethical codes of the culture. These qualities can be part of the super-ego or the internalized adult authority which holds the actions of the individual in check. The theoretic quality questions this authority and encourages the freedom of mind and spirit to which knowledge could lead.



In the study reported here, both educators and non-educators ranked the religious and ethical qualities above the theoretic on the ideal image. This could mean that adults feel that there should be some control over the actions of the adolescent. For them while the theoretic quality is important to the adolescent it should not supersede the control over adolescent actions which the ethical and religious qualities imply.

While students see the ethical and religious qualities as being of high importance to the adolescent they believe more than do adults that for the graduate a freedom of mind and spirit should prevail over those traits which serve to restrain him.

This notion of adult-imposed restraint as opposed to adolescent-desired freedom appears to be relevant to the conflict which exists between these two groups. It may well be that an extreme position in either direction could do much to retard the personal and social development of the adolescent in our society.

#### REFERENCES

1. BATTLE, H. J. Application of Inverted Analysis in a Study of the Relation Between Values and Achievement of High-School Pupils. Unpublished doctoral dissertation, The University of Chicago, 1954.
2. ERIKSON, E. H. *Childhood and Society*. New York: W. E. Norton, 1950.
3. GETZELS, J. W. Changing Values Challenge the Schools. *School Review*, Spring 1957, **65**, 97-102.
4. GILBERT, H. H. High school student opinions on reasons for failure in high school subjects. *J. Educ. Res.*, 1931, **23**, 46-49.
5. GOLDMAN, S. Sub-Public Perceptions of the High-School Graduate and the Roles of Institutions in His Development. Unpublished doctoral dissertation. The Univ. Chicago, 1961.
6. HESS, R. D., & GOLDBLATT, I. The status of adolescents in American society: a problem in social identity. *Child Devel.*, 1957, **28**, 459-468.
7. ———. Op. cit., 459-460.
8. JENKINS, D. H., & LIPPITT, R. Interpersonal perceptions in the classroom. In J. M. Seidman (Ed.), *The Adolescent: A Book of Readings*. New York: The Dryden Press, 1953. 583-599.
9. LEWIN, K. The Field Theory Approach to Adolescents. In J. M. Seidman (Ed.), *The Adolescent: A Book of Readings*. New York: The Dryden Press, 1953.
10. REMMERS, H. H., & WELTMAN, N. Attitude inter-relationships of youth, their parents and their teachers. *J. Soc. Psychol.*, 1947, **31**, 61-68.
11. WALLACE, D. L. Multiple Comparisons in the Analysis of Variance. The American Society for Quality Control. National Convention Transactions, 1957.

Department of Education  
Oklahoma State University  
Stillwater, Oklahoma

## A COMPARATIVE ANALYSIS OF OCCUPATIONAL INTEREST SCORES BETWEEN GIFTED AND TYPICAL 5TH GRADE PUPILS\*

*Lompoc Unified Schools, California*

---

RUSSELL N. CASSEL AND GENE HENDSCH

---

### A. PURPOSE

This study was concerned with making a comparative analysis between scores from two occupational interest inventories for 30 gifted and 34 typical 5th grade pupils from the same general school area. It sought to determine whether occupational interests as measured by the two psychological interest inventories for youth in the 5th grade classes were significantly different between gifted and typical individuals.

### B. INTEREST INVENTORIES

The Picture Interest Inventory—Grades 7 to Adult (5), and The Occupational Interest Inventory-Intermediate Series (4) were utilized. Both inventories were administered during the same setting, The Picture Interest Inventory usually being administered first. Pupils were told that the two inventories were being used experimentally for purpose of identifying the occupational interests of youth at the 5th grade level, and that scores on the inventories would be interpreted and explained to them at some later date.

### C. GROUPS

#### 1. *Gifted Pupils*

The gifted group included 30 pupils, about equally divided for sex, the total 5th grade gifted population in two elementary schools on the Vandenberg Air Force Base (AFB), California during the 1961-62 school year. All of these pupils had IQs of 130 or better on the basis of either The Wechsler Individual Scale for Children (WISC) or The California Test of Mental Maturity (CTMM) 1957 Short Form. Their IQs ranged from 130 to 160 with a mean ( $M$ ) IQ of 134.40, and with a standard deviation ( $SD$ ) of 10.66. They ranged in age from 121 to 132 months old with a  $M$  age of 128.00 months, and with a  $SD$  of 3.12 months. All pupils from both the

---

\* Received in the Editorial Office on May 16, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



gifted and typical groups were dependents of military personnel assigned to Vandenberg Air Force Base (AFB), California.

## 2. Typical Pupils

The typical group included 34 pupils, about equally divided for sex, the total membership of an average 5th grade class in one of the two elementary schools from which the gifted pupils were selected. The total mental scale IQ on the CTMM test referred to ranges from 85 to 127 with an (*M*) IQ of 110.40, and with an *SD* of 12.30. They ranged in age from 120 to 141 months with an *M* age of 128.34 months, and with an *SD* of 4.28 months.

TABLE 1  
DATA RELATIVE TO THE PICTURE INTEREST INVENTORY SCORES  
FOR GIFTED AND TYPICAL 5TH GRADE PUPILS  
(*N* = 64, 30 gifted and 34 typical)

Part scores	Validity $r_{pbis}$	Gifted group	Typical group	Differences	<i>t</i> - values
<i>Fields of Interest</i> (McCally <i>T</i> -scores; <i>M</i> = 50, and <i>SD</i> = 10)					
Inter-personal Ser.	.065				
<i>M</i>		70.10	73.79	3.69	0.53
<i>SD</i>		29.30	27.79	—1.51	0.30
Natural	.066				
<i>M</i>		41.37	44.74	3.37	0.53
<i>SD</i>		25.30	25.74	0.14	0.03
Mechanical	.055				
<i>M</i>		24.90	28.65	3.75	0.60
<i>SD</i>		31.17	36.12	4.95	0.84
Business	—0.040				
<i>M</i>		56.07	54.12	—1.95	0.32
<i>SD</i>		26.10	22.88	—3.22	0.75
Esthetic	.001				
<i>M</i>		62.24	62.29	0.05	0.07
<i>SD</i>		30.10	26.09	—4.01	0.80
Scientific	—0.036				
<i>M</i>		58.67	56.32	—2.35	0.29
<i>SD</i>		32.50	32.59	0.09	0.02
<i>Supplemental Scale</i>					
Verbal	—0.021				
<i>M</i>		64.47	63.32	—1.15	0.17
<i>SD</i>		25.60	28.84	3.24	0.69
Computational	.067				
<i>M</i>		46.97	50.59	3.62	0.54
<i>SD</i>		26.53	27.03	0.50	0.11
Time perspective	.090				
<i>M</i>		59.10	63.74	4.64	0.73
<i>SD</i>		27.07	24.09	—2.98	0.66

\*  $r_{pbis}$  of .250 significant at .05 level; # *t* of 2.00 significant at .05 level

\*\*  $r_{pbis}$  .325 significant at .01 level; ## *t* of 2.66 significant at .01 level

## D. RESULTS

1. *Picture Interest Inventory*

The discriminant function statistical technique as described by Horst (3) was utilized for making comparisons between the scores for the 30 gifted and 34 typical 5th grade pupils. From this discriminant function technique the multiple-point bi-serial  $R_{pbis\ c}$  corrected for shrinkage was  $.166 \pm .131$ . This would indicate that there were no significant differences between the gifted and typical scores on The Picture Interest Inventory. Table 1 contains the

TABLE 2  
DATA RELATIVE TO THE OCCUPATIONAL INTEREST INVENTORY SCORES  
FOR GIFTED AND TYPICAL 5TH GRADE PUPILS  
( $N = 64$ ; 30 gifted and 34 typical)

Part scores	Validity $r_{pbis}$	Gifted group	Typical group	Differences	<i>t</i> - values
<i>Fields of Interest</i> (McCally <i>T</i> -scores; $M = 50$ , and $SD = 10$ )					
Personal social	.001				
M		53.57	53.48	—0.09	0.02
SD		29.79	10.82	—18.97	4.74##
Natural	.155				
M		53.48	57.21	3.73	0.70
SD		10.83	28.90	18.07	4.88##
Mechanical	— .199				
M		40.39	27.62	—12.77	2.16#
SD		21.18	26.10	4.92	1.20
Business	.097				
M		35.75	40.00	4.25	0.79
SD		21.71	21.66	0.10	0.26
Art	.101				
M		51.75	56.21	4.46	0.81
SD		23.68	19.72	—3.96	1.04
Science	— .096				
M		53.93	48.93	—5.00	0.77
SD		23.93	27.76	3.83	0.85
<i>Types of Interest</i>					
Verbal	— .105				
M		42.57	37.93	—4.64	0.83
SD		24.21	20.03	—4.18	1.07
Manipulative	— .010				
M		41.50	41.10	—0.40	0.08
SD		20.29	21.24	0.95	0.26
Computational	.038				
M		49.68	51.38	1.70	0.30
SD		25.89	18.69	—7.20	1.80
<i>Level of Interest</i>					
M	.069	45.36	48.93	3.57	0.56
SD		22.57	28.78	6.21	1.38

\*  $r_{pbis}$  .250 significant at .05 level; # *t* of 2.00 significant at .05 level

\*\*  $r_{pbis}$  .325 significant at .01 level; ## *t* of 2.66 significant at .01 level



validity  $r_{pbis}$  indexes for the nine separate scores on this inventory from the discriminant function matrix, and on which the  $R_{pbis\ c}$  was obtained. Similarly, from Table 1 it should be observed that not a single one of the  $M$ s or  $SD$ s for the part scores obtained statistical significance between the gifted and typical pupils.

## 2. Occupational Interest Inventory

The discriminant function multiple-point bi-serial  $R_{pbis\ c}$  corrected for shrinkage was  $.252 \pm .128$ , and for ten variables this index is not statistically significant at the .05 level of confidence (2). This would indicate that there were no significant differences between the gifted and typical pupil scores on The Occupational Interest Inventory. Table 2 contains the validity  $r_{pbis}$  indexes for the ten separate scores on this inventory from the discriminant function matrix, and on which the  $R_{pbis\ c}$  was obtained. Similarly, it should be observed that not a single one of the  $M$ s for part scores obtained statistical significance, and only two of the  $SD$ s were significant.

## E. SUMMARY

This study was concerned with making a comparative analysis between scores from two occupational interest inventories (The Picture Interest Inventory, and The Occupational Interest Inventory) for 30 gifted and 34 typical 5th grade pupils. A discriminant function statistical technique after Horst was utilized for making the comparisons. No statistically significant differences were obtained for scores from either of the two interest inventories between gifted and typical pupils.

## REFERENCES

1. CASSEL, R. N. Comparing IBM and Hand Scoring Pad Administration of the Kuder Vocational Preference Record (publication pending).
2. GUILFORD, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1956.
3. HORST, P., & SMITH, S. The discrimination of two racial samples. *Psychometrika*, 1950, 15, No. 3.
4. LEE, E. A., & THORPE, L. P. Occupational Interest Inventory. The Calif. Test Bureau, Monterey, Calif., 1944.
5. WEINGARTEN, K. P. Picture Interest Inventory. The Calif. Test Bureau, Monterey, Calif., 1958.

*Lompoc Unified School District*  
*Lompoc, California*

## A STUDY OF THE EFFECTS OF MECHANICALLY INDUCED TENSION OF THE NECK MUSCLES ON THE PERCEPTION OF VERTICALITY\*

*Department of Psychology, Michigan State University*

CARL W. SCHNEIDER AND S. HOWARD BARTLEY

### A. PROBLEM

The present study stems from several considerations, some observational and some theoretical. The title ostensibly links neck tension as a condition for visual perception. Gibson (1) has pointed out that all visual responses we know about occur in a gravitational field, it inducing concurrent responses via other mechanisms such as kinesthesia which are not totally separate from vision. He believes up and down would have no meaning outside the gravitational field.

The ideas that Werner and Wapner (2) expressed in their sensory-tonic theory of perception are an example of connecting perceptual end-results and various body tension conditions. While we need neither describe their theory here in detail nor follow it in our experimental planning, their findings do demonstrate some active relationship between the structure of the seen field and muscular processes.

Tension is one of the effects produced in the muscles, either through their own contraction or through loads put upon them. The only mode of feedback from muscle action is through the kinesthetic sense organs which provide not only for effective posture or movement, but also self-awareness of position and motion. This self-awareness is not only kinesthetic but visual. Accordingly something may have a certain visual location because of some peculiar body position of the observer. At the same time body position may be experienced in a manner dependent upon the visual input to the central nervous system.

In the present study, a rotational torque was applied to the head while the observer maintained a fixed straight-ahead position. Thus a left-right asymmetry in neck-muscle tension is produced while the observer is in every respect in line with the true vertical. Under various induced tensional states of the neck muscles the observer is required to set a rod, rotatable around

\* Received in the Editorial Office on May 21, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



its midpoint in the frontal plane, so that it appears to be in the true vertical.

Specifically, the present study was designed to investigate the effect of increasing neck-muscle tension on the perception of verticality and the effect of initial rod position on the adjustment of the rod to the vertical position.

## B. METHOD AND PROCEDURE

### 1. *Stimulus Object*

The visual target was an illuminated white rod, 24 inches long and one inch in diameter, which revolved around its midpoint in the frontal plane in an illuminated room. The target was 18 feet from the observer. Each observer was required to set the rod from some initial position to what he perceived as true vertical. Its position could be read on a protractor to an accuracy of  $0.25^\circ$ .

### 2. *Apparatus*

The apparatus consisted mainly of a revolving disk,  $15\frac{1}{2}$  inches in diameter, which turned  $30^\circ$  clockwise in the horizontal plane when weights were applied to it. Braces projecting downward from the center of the disk held the observer's head. A rotational tension was placed on the neck when the disk was loaded with a weight.

### 3. *Procedure*

Four different amounts of tension were used in the experiment: 0 (control—no tension), 1 (weight of gear strapped on), 2 (2.7 Kg weight applied), 3 (4.4 Kg weight applied). There were seven initial rod positions:  $0^\circ$ ,  $+5^\circ$ ,  $-5^\circ$ ,  $+30^\circ$ ,  $-30^\circ$ ,  $+90^\circ$ ,  $-90^\circ$ . All of the initial rod positions were combined with the four degrees of tension to make a total of 28 different conditions. The 28 conditions were randomized, and each observer performed one setting under each condition. The observer had 20 seconds to make his setting, and there were 30 seconds between each setting.

Thirty-five male students from an introductory course in psychology served as observers.

### 4. *Measures*

The measures were degrees of deviation of the observer's setting from true (plumb line) vertical. Clockwise deviations from the true vertical were designated as (+) positive, and counterclockwise ones (—) negative.

## C. RESULTS

Table 1 shows the algebraic means for the 35 observers under each of the 28 conditions. The control (no tension) is  $+107^\circ$ . The rod is perceived

TABLE 1  
ALGEBRAIC MEAN POSITION OF THE PERCEIVED VERTICAL WITH REFERENCE  
TO THE TRUE VERTICAL FOR ALL OF THE TWENTY-EGHT CONDITIONS

Initial rod position	Tension				Grand mean
	0	1	2	3	
0°	.021	.035	.078	.071	.051
+5°	.257	.385	.978	.900	.630
-5°	-.242	-.392	.171	-.107	-.142
+30°	.050	.321	.428	.650	.362
-30°	.200	.142	.114	.228	.171
+90°	.292	.307	.657	.828	.521
-90°	.171	.235	.207	.485	.490
Grand mean	.107	.147	.376	.436	

as vertical, on the average, when it is rotated  $.107^\circ$  clockwise. The grand means for the four degrees of tension indicate the perceived vertical tends to deviate farther from the true vertical as tension is increased. However, the grand means for the various initial rod positions show a different type of relationship. The perceived vertical for the positive initial rod deviates farther from the true vertical than that for the negative initial rod positions. The grand means for the  $+5^\circ$ ,  $+30^\circ$  and  $+90^\circ$  are larger than those for  $-5^\circ$ ,  $-30^\circ$  and  $-90^\circ$  respectively. This would seem to indicate that the initial rod position did influence the perceived vertical.

The data from the experiment were analyzed by a triple classification analysis of variance. It was found that the effect of tension was significant,  $F = 15.89$ ,  $df = 3$ , and 102,  $p .01$ . The interaction between tension and observers was significant,  $F = 1.54$ ,  $df = 102$ , and 612,  $p .01$ .

The effect of initial rod position was also significant,  $F = 16.10$ ,  $df = 6$ , and 204,  $p .01$ . Interaction between observers and initial rod position was significant,  $F = 1.65$ ,  $df = 204$ , and 612,  $p .01$ .

Interaction between tension and initial rod position was not significant.

Twenty-seven of the observers reported that the rod appeared "fuzzy" or blurred under the neck tensional conditions. Three of the observers reported that the rod seemed to fade away. One observer reported that the rod disappeared for a few seconds.

#### D. FURTHER CHECK ON CERTAIN RESULTS

The reports of blurring induced us to ask an optometrist to try and determine the cause.

Retinoscopy was performed on two observers fixating on a penlight 18 feet away. The two conditions of maximal tension produced a shift in accommo-



dation toward hyperopia for the weights of  $+0.25$  and  $+0.50$  diopters, respectively. This shift lasted only while the weights were in use.

### E. DISCUSSION

Neck-muscle tension does affect the perceived vertical, the amount depending upon the amount of tension. The perceived vertical under tensional states 1, 2 and 3 is clockwise of the control (no tension) position and also in the direction opposite of the side of increased neck tension.

The deviation of the observer's setting from true vertical was smaller when the rod approached the vertical from the negative position than when it approached from the positive position. This was most marked when the initial position was  $+5^\circ$  and  $-5^\circ$ . This may have been due to the fact that many of the observers tended initially to accept these rod positions as vertical, particularly under tensional states 2 and 3.

An important factor possibly contributing to the differences in the mean perceived vertical of the negative and positive initial position is the fact that an observer in making a setting from a negative initial position, must allow the rod to travel a greater distance to reach a point near the mean perceived vertical for the particular tensional state involved. The observer may become cautious as the rod nears the true vertical and tend to stop the rod short of true vertical at times.

### F. SUMMARY

A connection between kinesthesia and visual perception was demonstrated. As tension in the neck muscles produced by a horizontal torque is increased, the perceived vertical of an upright rod tends to deviate farther from the true gravitational vertical on the side opposite the application of the torque.

The initial rod position had a significant effect on the observer's perceived vertical.

Hyperopia tends to result when torque becomes great, possibly accounting for the blurred vision experienced by many of the observers under the same conditions.

### REFERENCES

1. GIBSON, J. J. Perception of the Visual World. New York: Houghton Mifflin, 1950.
2. WERNER, H., & WAPNER, S. Sensory-tonic theory of perception. *J. Pers.*, 1949, 18, 88-107.

*Department of Psychology*  
*Michigan State University*  
*East Lansing, Michigan*

\$20.00 per annum  
\$15.00 per volume  
\$7.50 per half volume

QUARTERLY  
Two volumes per year  
Immediate publication

October, 1962  
Volume 54, Second Half

(Founded in 1915 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

OCTOBER, 1962

(Manuscripts are printed in the order of final acceptance)

Effects of success and failure and the competence of the model on the acquisition and reversal of matching behavior . . . . .	251
BY MILTON E. ROSENBAUM, DOUGLAS K. CHALMERS, AND WILLIAM C. HORNE	
New leader behavior description subscales . . . . .	259
BY RALPH M. STODGILL, OMAR S. GOODE, AND DAVID R. DAY	
Formation and evaluation of hypotheses in dream interpretation . . . . .	271
BY ERIKA FROMM AND THOMAS M. FRENCH	
Patient wisdom: An anthology of creative insights in psychotherapy . . . . .	285
BY STEWART B. SHAPIRO	
Perceptual functioning in the duckling . . . . .	293
BY NICHOLAS PASTORE	

(OVER)

Copyright, 1962, by The Journal Press  
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at  
Provincetown, Mass., under the Act of March 3, 1879  
Second-class postage paid at Provincetown, Mass.



Reaction-time as a function of stimulus intensity among normal and psychotic subjects . . . . .	299
BY H. E. KING	
Predictability of multidimensional absolute identifications from information transmitted with unidimensional stimuli . . . . .	309
BY GLENN R. HAWKES	
Increased occurrence of EEG alpha during increased attention . . . . .	317
BY THOMAS MULHOLLAND AND SYLVIA RUNNALS	
Personality and attitudinal correlates of response to drug treatment in psychiatric outpatients: I. Theoretical orientation and measuring instruments to be employed . . . . .	331
BY K. RICKELS AND R. W. DOWNING	
Personality and attitudinal correlates of response to drug treatment in psychiatric outpatients: II. Preliminary investigation and implications for further research . . . . .	345
BY R. W. DOWNING AND K. RICKELS	
"Hypnotic behavior" as a function of task motivation . . . . .	363
BY THEODORE XENOPHON BARBER AND DAVID SMITH CALVERLEY	
Predicting organization in changed communication networks . . . . .	391
BY ARTHUR M. COHEN AND WARREN G. BENNIS	
Measurement of the primary factors of the total personality . . . . .	417
BY THOMAS N. JENKINS	
Differential impairment in multiple sclerosis . . . . .	443
BY CHARLES A. KNEHR	
A discussion of chronogeometry as related to non-stationary time series such as the physiological electrogram . . . . .	453
BY SAMUEL J. M. ENGLAND AND BENJAMIN PASAMANICK	
Factor analyses of schizophrenic and organic test data . . . . .	467
BY CHARLES A. KNEHR	
The continuity-discontinuity controversy in paired-associate learning . . . . .	473
BY NELSON BUTTERS AND MORTON WIENER	
Consistency between attitudinal affect and spontaneous cognitions . . . . .	485
BY MILTON J. ROSENBERG AND PHILIP K. OLTMAN	
Stimulus size and Rorschach responses . . . . .	491
BY IJA N. KORNER, ROGER B. ALLISON, JR., AND MAX D. ZWANZIGER	

## EFFECTS OF SUCCESS AND FAILURE AND THE COMPETENCE OF THE MODEL ON THE ACQUISITION AND REVERSAL OF MATCHING BEHAVIOR\*<sup>1</sup>

*Department of Psychology, State University of Iowa*

---

MILTON E. ROSENBAUM, DOUGLAS K. CHALMERS, AND WILLIAM C. HORNE

---

### A. INTRODUCTION

A number of recent laboratory studies have examined variables that affect the learning of imitative behavior. Two of these variables are treated in the present experiment in order to evaluate the possible interactive effects of their simultaneous manipulation on acquisition and reversal of an imitative response. Available evidence indicates that the degree of an individual's competence in a task prior to the availability of a model affects his imitative performance upon the introduction of a model (4, 7). Kanareff and Lanzetta (2) have demonstrated that the experience of failure as opposed to success in a prior task leads to facilitation of the acquisition of a matching response in a criterion task.

Another variable of interest is the competence of the model for imitative behavior (8). Rosenbaum and Tucker (12) have shown that the greater the apparent competence of the model, the greater is the facilitation of learning of matching behavior.

These two variables are considered in the present instance because of the possibility that when presented simultaneously the effects of one variable may amplify or obscure the contribution of the other. For example, failure may result in a disposition to match such that the competence of the model may be rendered irrelevant to the learning process. On the other hand, the presentation of a highly competent model may overcome any contribution of the success or failure of the performer in a previous task.

Information concerning the extinction or transfer of social responses is meager and somewhat inconsistent (2, 3, 11, 13, 14). In the present ex-

---

\* Received in the Editorial Office on May 2, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported in part by a grant to the senior author from the National Science Foundation (G-16011). The data were collected while the junior authors were participants in the National Science Foundation Undergraduate Research Participation Program.



periment in which *Ss* learn a discriminative response, a reversal of the probability of reinforcement for imitative behavior was introduced following the acquisition period.

### B. METHOD

The apparatus was identical to that employed in other experiments (12). Each of two *Ss* in each experimental session sat in separate rooms facing a panel on which signal lamps designated the onset of a trial for *S* or his partner, "A." Also presented on the panel was an indication of the responses of A and *S* and, by means of jeweled lamps, the accuracy of the responses. Each *S* was informed that he was "B" and was thereby led to believe that the other *S* was A. Actually A's responses were simulated from a central control panel operated by *E*.

Each *S* was told that on signal he was to predict on each trial the outcome of a facsimile horse race by indicating whether his horse, running against a field of horses, would win or lose the race. In an initial series of 20 trials, *S* presented his predictions alone and during a corresponding interval of time A was ostensibly presenting similar responses. On each of the 20 trials *S* was informed of the accuracy of his response. On the basis of a prearranged schedule, *Ss* assigned to the Success condition were designated correct on 80 per cent of the trials while Failure *Ss* were arbitrarily correct on only 20 per cent of these trials.

Instructions were then given for the second series of trials (the acquisition series) during which *Ss* were to respond in sequence; A first, and B (*S*) second. A's and B's horses were to be thought of as running on separate tracks and as not competing against each other. On each trial, *E* presented A's predictions on *S*'s panel. Half of A's predictions were win and half lose on a random basis. The *E*, employing programmed schedules, also designated the accuracy of A's responses. For half of the pairs of *Ss*, A's responses were correct on eight of each block of 10 trials to create the High Model's Competence (HMC) condition. For the remaining *Ss*, who were in the Low Model's Competence (LMC) condition, A was correct on only two of each block of 10 trials.

The *S*'s responses on 80 per cent of the trials were designated as correct if *S* matched the response of A and arbitrarily, but according to a prearranged schedule, *S* was incorrect on two of every 10 trials; if he matched A's response on those trials. The second series included 100 trials.

During the final series of 80 trials, the reversal series, *S* was correct on eight of each block of 10 trials if he nonmatched A's responses and correct on

a randomly selected two trials of 10 if he matched the response of A. The reversal series was initiated without warning to S.

Sixty-four students in the introductory psychology course were recruited in pairs of the same sex and were assigned at random to four experimental conditions created by the Success and Failure conditions and the two Model's Competence conditions. Eight males and eight females participated in each condition.

### C. RESULTS

#### 1. Acquisition

The results for the acquisition of the imitative response indicates confirmation of previous findings. The curves in Figure 1 which present the mean frequency of matched responses on each block of 10 trials indicate that Ss who had experienced failure learned on the average to a superior level as compared to those Ss who had experienced success (2). The main effect for

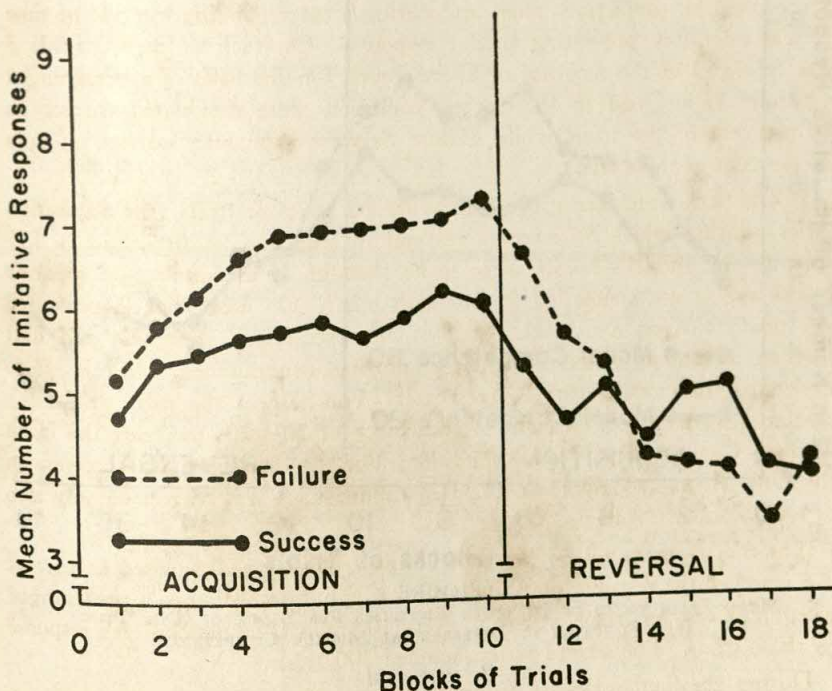


FIGURE 1  
MEAN FREQUENCIES OF MATCHED RESPONSES PER BLOCK OF TEN TRIALS  
AS A FUNCTION OF PRIOR SUCCESS AND FAILURE



Prior Experience is significant beyond the .01 level ( $F = 10.43$ ,  $df$  1/56) by an extended Type III analysis of variance (6).

Figure 2 presents the results for the Model's Competence variation. During the acquisition period, Ss in the HMC condition learned to imitate on the average to a higher level than Ss in the LMC condition (12). In addition to a significant main effect for Model's Competence ( $F = 5.79$ ,  $df$  1/56), a significant Trials  $\times$  Model's Competence interaction ( $F = 2.10$ ,  $df$  9/504) suggests that the difference between the Model's Competence groups developed over the acquisition series. The main effect for Sex and other interactions of the experimental variables with each other and with Trials were not significant but the main effect for Trials was significant beyond the .01 level ( $F = 8.54$ ,  $df$  9/504).

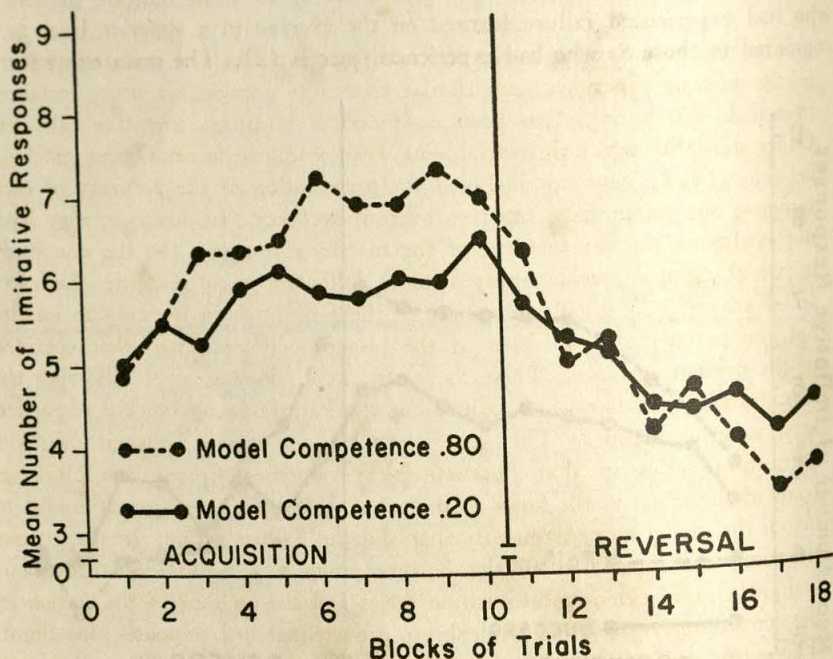


FIGURE 2  
MEAN FREQUENCIES OF MATCHED RESPONSES PER BLOCK OF TEN TRIALS  
AS A FUNCTION OF DIFFERENTIAL MODEL'S COMPETENCE

## 2. Reversal

During the initial trials of the reversal procedure, Failure Ss continued to make a greater mean frequency of matched responses than did Success Ss.

By the fourth block of reversal trials, however, this relationship was reversed. Only in the final block of trials did the Failure Ss again present more frequent matched responses (see Figure 1). Analysis of variance of the reversal trials indicates in addition to a significant main effect for Trials ( $F = 16.21$ ,  $df = 7/392$ ), a significant Trials  $\times$  Prior Experience interaction ( $F = 5.86$ ,  $df = 7/392$ ).

After the initial block of reversal trials in which Ss in the HMC condition match more frequently than LMC Ss, fluctuations occur between the two conditions until in the last three blocks of trials the HMC group consistently presents fewer matched responses (Figure 2). A significant Trials  $\times$  Model's Competence interaction ( $F = 2.76$ ,  $df = 7/392$ ) reflects these trends. All other main effects and interactions are not significant.

#### D. DISCUSSION

The results pertaining to the effect of prior experience of success or failure in the present experiment are similar to results obtained in other instances in which this variable has been examined as it affects imitative behavior. Other studies in which this variable has been examined in relation to matching behavior (4, 7) have not included the presentation of the accuracy of each response but rather have reported over-all accuracy. In addition, they have not evaluated the development of the matching response. In the one study in which accuracy was reported to S following each response, Lanzetta and Kanareff (2) found superiority of the Failure Ss in learning to imitate but the pattern of acquisition of the response differed from that obtained in the present instance. Their Ss started at a common level following the manipulation of success and failure but the Failure group quickly responded in a superior manner. This superiority did not remain constant over the acquisition series in that fluctuations in response affected one block of trials in the midst of the series to the extent that the Success group tended to match the model more frequently than did the Failure group. In the present situation, the superiority of the Failure group was maintained throughout a slightly more extensive acquisition series and the amount of fluctuation in the continuously ascending tendency to present matched responses was slight. The major difference in the procedures of the two studies appears to be that Kanareff and Lanzetta employed as the pre-training task one that was similar but not identical to the one used for training of the imitative response. In the present study, the pretraining task was essentially identical to that presented for the training series. Generalization of the knowledge of personal competence was to a situation more distant from the original source in



Kanareff and Lanzetta's experiment and may have led to more disturbance in the application of the obtained knowledge of competence than occurred in the present study.

The results obtained for variation of the model's competence are consistent with the previous findings of Rosenbaum and Tucker (12) in indicating superior acquisition of the matching responses in the presence of a more competent model. No interaction of the Prior Experience variations with the Model's Competence treatments is indicated.<sup>a</sup> This finding may be contrasted with the findings of Mausner and Bloch (9) who examined matching behavior without report on accuracy in relation to the simultaneous presentation of the two independent variables. They found that Failure *Ss* conformed more than Success *Ss* in the presence of a competent partner but no difference is indicated when a relatively incompetent model is employed. An important difference in procedure between the two experiments is that Mausner and Bloch revealed the competence of the model prior to the series in which conformity was assessed while in the present experiment competence of the model was revealed slowly through the second series.

In interpretation of the present findings, the earlier experience of failure is likely to have led to lack of confidence on the part of *S* in his capabilities to perform the required task, and he subsequently employed the cues from a model to a greater extent than did an *S* who had been previously successful. The tendency to orient toward another in the present situation which is likely to have been rewarded in other failure situations provided an advantage in acquiring the appropriate imitative response. This advantage appeared from the start and was maintained throughout the training series.

In contrast, the variations in Model's Competence produced no difference in early trials but led finally to superior performance in the HMC condition. As Rosenbaum and Tucker (12) have indicated, despite the irrelevancy of the information concerning the accuracy of the model for successful task performance, presentation of a more competent model served to facilitate learning. This may be interpreted as reflecting earlier imitative behavior wherein competent models have been found to provide useful behavioral cues. In the present situation, knowledge as to competence of the model is gained gradually during the acquisition period and therefore may be a less potent influence on response than when competence is stated beforehand and less subtly.

The data for the reversal period indicate that the conditions leading to superior learning in the acquisition period also are related to superior performance in responding to the reversed discrimination.

In one view (e.g., 10) this result was unexpected. In order to learn the reversed discrimination, extinction of the originally acquired response should be necessary. For Ss who had learned to a superior level, and received more frequent reinforcement, the extinction process should be delayed.

In several recent studies of reversal training results have been obtained that are in some respects similar to present findings. Reversal of a discrimination is learned more rapidly following reinforced overtraining of the correct response (e.g., 1). These results have been interpreted as based on the earlier recognition of change in the reinforcement pattern. In the present experiment the groups that had performed to a superior level during the acquisition period had been performing the correct response more frequently and had thereby received a greater number of reinforcements prior to the introduction of the reversal procedure.

Recent studies of concept learning suggest an alternative interpretation. In addition to acquisition of the specific imitative response during the training series, it is likely that a variety of other responses were being performed by Ss in this experiment. Among these responses one can conceive a variety of mediating responses including orienting responses (15), and conceptual responses (5). As has been suggested earlier, the prior experience of failure may lead to a generalized learned response of orienting toward the behavior of others. Similarly, prior experience in many social situations may produce ready orientation to competent models when cues indicating competency of another are present.

In both training periods, orienting responses of this character are likely to facilitate the required response, be this matching or non-matching behavior and in addition a conceptual dimension of relevance of the behavior of another may be formed.

To test these alternative interpretations, Ss in a Success group and in an LMC condition should receive additional reinforced trials comparable to the number received by their counterparts in the remaining conditions. If the reversal results after equating acquisition levels were similar to those obtained in the present experiment, the interpretation involving orienting responses and concept formation would be supported.

#### E. SUMMARY

Differential personal competence on a prior task and variations in the competence of a model for imitation were included as independent variables in a factorial design in which acquisition and reversal of an imitative response were examined. The results for the acquisition period, as in previous experi-



ments, showed that prior failure in contrast to success led to facilitation of the imitative response and that the more competent the model, the greater is the tendency to match his responses. No interaction of these variables was indicated. Similarly, the results for a succeeding series of trials in which the probability of reinforcement was reversed indicated facilitation of non-imitation for Ss who experienced prior failure and who were presented with a high competence model. Alternate interpretations of the results were offered.

## REFERENCES

1. CAPALDI, E. J., & STEVENSON, H. W. Response reversal following different amounts of training. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 195-198.
2. KANAREFF, V. T., & LANZETTA, J. T. Effects of success-failure experiences and probability of reinforcement upon the acquisition and extinction of an imitative response. *Psychol. Rep.*, 1960, **7**, 151-166. (a)
3. ———. Effects of task definition and probability of reinforcement upon the acquisition and extinction of imitative responses. *J. Exp. Psychol.*, 1960, **60**, 340-348. (b)
4. KELMAN, H. C. Effects of success and failure on "suggestibility" in the auto-kinetic situation. *J. Abnorm. & Soc. Psychol.*, 1950, **45**, 267-285.
5. KENDLER, H. H., & KENDLER, T. S. Vertical and horizontal processes in problem solving. *Psychol. Rev.*, 1962, **69**, 1-16.
6. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.
7. MAUSNER, B. The effect of prior reinforcement on the interaction of observer pairs. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 65-68.
8. ———. The effect of one partner's success or failure in a relevant task on the interaction of observer pairs. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 557-560.
9. MAUSNER, B., & BLOCH, B. L. A study of the additivity of variables affecting social interaction. *J. Abnorm. & Soc. Psychol.*, 1957, **54**, 250-256.
10. McCULLOCH, T. L., & PRATT, J. G. A study of the pre-solution period in weight discrimination by white rats. *J. Comp. Psychol.*, 1934, **18**, 271-290.
11. MILLER, N. E., & DOLLARD, J. Social Learning and Imitation. New Haven: Yale Univ. Press, 1941.
12. ROSENBAUM, M. E., & TUCKER, I. F. The competence of the model and the learning of imitation and non-imitation. *J. Exp. Psychol.*, 1962, **63**, 183-190.
13. SCHEIN, E. H. The effect of reward on adult imitative behavior. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 389-395.
14. SOLOMON, R. L., & COLES, M. R. A case of failure of generalization of imitation across drives and across situations. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 7-13.
15. WYCKOFF, L. B., JR. The role of observing responses in discrimination learning. Part I. *Psychol. Rev.*, 1952, **59**, 431-442.

Department of Psychology  
State University of Iowa  
Iowa City, Iowa

## NEW LEADER BEHAVIOR DESCRIPTION SUBSCALES\*

*Ohio State University and Carnegie Institute of Technology*

---

RALPH M. STOGDILL, OMAR S. GOODE, AND DAVID R. DAY

---

### A. INTRODUCTION

The research reported in this paper is part of a study designed to isolate, if possible, dimensions of observable behavior that appear to be of theoretical importance to a study of leadership. Various factorial studies of leader behavior have been reviewed by Carter (3), Borgatta, Cottrell and Meyer (2), Stogdill (9), and Bass (1).

The present research is based on a modification and expansion of the Leader Behavior Description Questionnaire, which grew out of work initiated by Hemphill (6). The development of the questionnaire has been described by Hemphill and Coons (7). Shartle (8) has outlined the theoretical considerations underlying the descriptive method. Factor analyses of the items, particularly by Halpin and Winer (5) and Fleishman (4), have consistently yielded two strongly defined factors, identified as *Consideration* and *Initiation of Structure*. Items on the Consideration scale describe behaviors that indicate a regard for the comfort and well being of the group members, as well as an equalitarian respect for the members' contributions. Items in the Initiation of Structure scale describe behaviors that clearly define the leader's role, and structure the expectations of the members.

Recent theoretical work by Stogdill (9) on the differentiation of positions and roles in organized groups suggested that two factors are insufficient to account for the observable variance in leader behavior. Working from theoretical considerations and a survey of the literature, items were developed for the following hypothetical subscales, in addition to Consideration and Initiation of Structure.

- Representation of member interests (5 items)
- Tolerance of uncertainty (10 items)
- Persuasion (10 items)
- Retention of the leadership role (10 items)
- Tolerance of member freedom of action (10 items)
- Production emphasis (10 items)

---

\* Received in the Editorial Office on May 22, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



Predictive accuracy (10 items)

Reconciliation of conflicting demands (5 items)

Each item in the questionnaire is answered by circling one of five possible responses: always, often, occasionally, seldom, never. Each item receives a score that can range from 5 to 1 (or 1 to 5 for negatively stated items), as shown in the following examples.

He is very persuasive talker

Always (5) Often (4) Occasionally (3) Seldom (2) Never (1)

He fails to take necessary action

Always (1) Often (2) Occasionally (3) Seldom (4) Never (5)

### B. THE POPULATIONS SAMPLED

Descriptions were obtained of two samples of subjects: (1) ministers of various religious denominations and (2) leaders in community development activities. For descriptions of community leaders, one questionnaire was mailed to each of 150 citizens throughout Ohio who had participated in a Community Development Conference at The Ohio State University. Each recipient of a questionnaire was asked to describe the leader behavior of "a citizen who is active in community development activities and whom you know well enough to describe accurately as a leader in community development." Of the 57 usable questionnaires returned, 25 described business men, 18 professional men, 12 city and county officials, and 2 women's club presidents. The communities ranged in size from approximately one thousand to more than a million.

For the ministers, two questionnaires each were mailed to 150 of the approximately 600 ministers located in Columbus, Ohio, and vicinity. Each recipient was requested to "give the enclosed questionnaires to two members of your congregation who know you well enough to describe you accurately as a leader of your congregation." The 103 usable returns were found to be representative of the various Protestant, Catholic and Jewish denominations, with the following exceptions. United Presbyterian churches are over-represented. Baptist, Community Nazarene, Pentacostal Apostolic, and Spiritualist Churches are under-represented. It should be kept in mind that the 103 questionnaires do not represent 103 separate ministers, since some were probably described twice.

### C. METHOD OF ANALYSIS

The subscores were intercorrelated and factor analyzed. The method of principal components was used for the factor analysis. In this method, ones

are entered in the diagonals, and all the variance of a test is completely accounted for by the factor loadings for that test. That is, all commonalities should be 1.00. The number of factors equals the number of tests. All except the general factor tend to be bipolar, with about half the loadings positive and the other half negative.

In rotating the factors, an effort was made to search for specific factors rather than to account for the maximum variance in the smallest possible number of factors. The purpose of this procedure was to determine which subscales exhibit the greatest likelihood of independence from the general and subgeneral factors. If, under this procedure, a subscale fails to exhibit any appreciable loading as a specific factor, the hope of increasing its independence by the revision or addition of items is reduced, particularly for the population being studied.

#### D. RESULTS: COMMUNITY LEADERS

The subscale intercorrelations for community leaders are shown in Table 1. The results of the factor analysis are shown in Table 2.

Factor I, a general factor, has its high loadings on Persuasiveness, Initiating Structure, Role Enactment, Predictive Accuracy, Production Emphasis, and Demand Reconciliation; with a moderately high loading (.58) on Consideration. All the subscales except Consideration tend to describe behaviors that initiate structure; maintain role dominance; and maintain order, system, and goal direction. The factor might be called the *Initiation and Maintenance of Structure*.

Factor II has its highest loading (.92) on Tolerance of Uncertainty. There is a loading of .33 on Consideration and —.45 on Production Emphasis. This factor is essentially a measure of *Tolerance of Uncertainty*.

Factors III, IV and V are specific factors. Factor III has its only high loading (.93) on *Representation of Member Interests*. Factor IV is a measure of Tolerance of *Member Freedom of Action*, while Factor V describes *Consideration*.

Factor VI shows loading of —.34 on Role Enactment, —.33 on Reconciliation, and .46 on Persuasiveness. Contrary to what is found in Factor I, Persuasiveness is negatively related to role enactment. In community development activities, it would appear that the leader who resorts to persuasiveness is perceived as tending to relinquish the leadership role to other persons. The factor might be called *Attempt to Retain the Leadership Role through Persuasion*.

In Factor VII, Demand Reconciliation shows a loading of .33, and Role



TABLE 1  
INTERCORRELATION OF TEN SUBSCALES FOR FIFTY-SEVEN COMMUNITY LEADERS; RESIDUALS\* BELOW THE DIAGONAL

Subscale	1	2	3	4	5	6	7	8	9	10
1. Representation		.01	.36	.23	.36	.22	.20	.00	.23	.30
2. T. uncertainty	.00		.03	-.21	-.19	.26	.06	.40	-.47	-.01
3. Persuasiveness	.00	.00		.65	.56	.36	.73	.08	.58	.54
4. Structure	.00	.00	.00		.70	.46	.72	.05	.73	.68
5. Role enactment	.00	.00	.00	.00		.35	.62	-.13	.65	.74
6. Consideration	.00	.00	.00	.00	.00		.48	.40	.28	.52
7. Predictive acc.	.00	.00	-.01	.00	.00	.00		.10	.57	.70
8. T. freedom	.00	.00	.00	.00	.00	.00	.00		-.15	-.02
9. Production	.00	.00	.00	.00	.00	.00	.00	.00		.54
10. Reconciliation	.00	.00	.00	.00	.00	.00	.00	.00	.00	
Mean	19.6	37.7	39.5	37.2	39.8	41.1	39.5	18.2	17.7	19.7
Standard deviation	2.4	5.6	5.5	5.7	5.6	4.7	4.9	2.5	3.4	3.3
Reliability	.59	.85	.79	.72	.83	.77	.62	.86	.79	.58

\* Residuals on the diagonal are all .00, except -.01 in Column 4

TABLE 2  
FACTOR LOADINGS FOR COMMUNITY LEADERS<sup>a</sup>

Subscale	Factor									
	I <sub>1</sub>	II <sub>1</sub>	III <sub>1</sub>	IV <sub>2</sub>	V <sub>2</sub>	VI <sub>1</sub>	VII <sub>2</sub>	VIII <sub>1</sub>	IX <sub>1</sub>	X <sub>2</sub>
1. Representation	.22	-.12	.93	.20	.09	.00	.09	.07	-.01	.03
2. T. uncertainty	-.16	.92	.12	.21	.01	.06	-.10	.16	.16	-.11
3. Persuasiveness	.76	.11	.25	.10	-.23	.46	.06	-.17	-.10	-.18
4. Structure	.90	-.11	-.01	.04	-.03	-.01	-.05	.36	-.20	-.06
5. Role enactment	.80	-.06	.27	-.16	-.16	-.34	-.28	-.15	.01	-.10
6. Consideration	.58	.33	.02	.27	.68	-.04	-.08	-.09	-.04	-.03
7. Predictive acc.	.86	.26	.02	-.02	-.13	.17	-.01	-.03	-.01	.37
8. T. freedom	.07	.28	-.19	.93	.02	-.12	.05	-.05	.03	.02
9. Production	.80	-.45	.02	-.08	.00	.13	-.02	.06	.35	-.07
10. Reconciliation	.82	.20	.18	-.17	.04	-.33	.33	-.04	.03	-.01
% Variance	44.9	13.8	10.8	10.9	5.7	5.0	2.2	2.3	2.0	2.0



Enactment a loading of  $-.28$ . In view of the impersonal character of the behavior suggested by this factor, it might be called *Non-interpersonal Control*.

Factors VIII, IX, and X are weak specific factors. Factor VIII shows a loading of  $.36$  on *Initiating Structure*, while IX has its only appreciable loading ( $.35$ ) on *Production Emphasis*. Factor X appears with a weighting of  $.37$  on *Predictive Accuracy*.

Despite a strong general factor which accounts for 45 per cent of the total factor variance, several of the subscales exhibit sufficiently high loadings on specific factors to suggest that they may have some value as measures of discrete aspects of leader behavior in community development activities. The most promising of these subscales are Representation, Tolerance of Member Freedom of Action, and Consideration. Three other subscales, with only moderately high loadings on specific factors, are Initiating Structure, Production Emphasis and Predictive Accuracy.

#### E. RESULTS: MINISTERS AS LEADERS

The subscale intercorrelations and factor loadings for ministers are shown in Tables 3 and 4.

Factor I is a general factor, with its highest loadings on Initiating Structure, Role Enactment, Persuasiveness, Representation, Production Emphasis and Reconciliation of Conflicting Demands. This is essentially a measure of structuring behaviors. The factor might be called the *Initiation and Maintenance of Structure*.

Factor II is a subgeneral factor, with high loadings on Tolerance of Uncertainty, Tolerance of Member Freedom of Action, Consideration, and Demand Reconciliation. This factor is descriptive of behaviors that remove restraints to group action and provide scope for activities. It might be given the name *Facilitation of Group Action*.

Factors III and IV are specific factors. Factor III has its only high loading ( $.74$ ) on *Predictive Accuracy*. Factor IV shows a high weighting ( $.61$ ) on *Representation of Member Interests*.

Factor V has its highest loading ( $.67$ ) on Production Emphasis, with a moderately high weighting ( $.30$ ) on Tolerance of Freedom of Action. The factor is identified as *Motivation of Productive Activities*.

Factor VI has a loading of  $.50$  on Persuasiveness and a weighting of  $-.43$  on Role Enactment. This factor is identified as *Attempt to Retain Leadership through Persuasion*.

Factor VII has a loading of  $.47$  in Initiating Structure and a moderate

TABLE 3  
INTERCORRELATION OF SUBSCALES FOR 103 DESCRIPTIONS OF MINISTERS; RESIDUALS\* BELOW THE DIAGONAL

Subscale	1	2	3	4	5	6	7	8	9	10
1. Representation		.08	.43	.45	.01	.52	.33	.23	.19	.33
2. T. uncertainty	.00		.13	— .10	.59	.06	— .23	.52	.20	.43
3. Persuasiveness	.00	.00		.62	.20	.45	.47	.57	.45	.52
4. Structure	.00	.00	.00		— .04	.61	.44	.33	.38	.41
5. T. freedom	.00	.00	.00	.00		— .14	.01	.49	.36	.31
6. Role enactment	.00	.00	.00	.00	.00		.30	.40	.22	.54
7. Production	.00	.00	.00	.00	.00	.00		.18	.34	.19
8. Consideration	.00	.00	.00	.00	.00	.00	.00		.43	.60
9. Predictive acc.	.00	.00	.00	.00	.00	.00	.00	.00		.35
10. Reconciliation	.00	.00	.00	.00	.00	.00	.00	.00	.00	
Mean	20.4	37.5	42.1	38.7	37.5	41.5	34.9	42.5	41.0	19.8
Standard deviation	2.4	6.3	4.7	4.9	6.0	5.4	5.1	5.8	4.6	3.1
Reliability	.55	.84	.77	.70	.75	.75	.59	.85	.83	.77

\* Residuals on the diagonal are all .00



TABLE 4  
FACTOR LOADINGS FOR MINISTERS

Subscale	I <sub>1</sub>	II <sub>1</sub>	III <sub>1</sub>	IV <sub>1</sub>	V <sub>2</sub>	VI <sub>3</sub>	VII <sub>1</sub>	VIII <sub>2</sub>	IX <sub>1</sub>	X <sub>1</sub>
1. Representation	.66	.07	— .05	.61	— .22	— .26	— .17	.13	— .10	— .08
2. T. uncertainty	— .17	.85	— .15	.10	— .13	— .18	— .07	.03	.40	.00
3. Persuasiveness	.75	.36	— .02	.10	— .05	.50	— .01	— .10	.06	.19
4. Structure	.84	.04	.09	— .06	— .09	.01	.47	— .10	.13	— .16
5. T. freedom	— .14	.82	.17	.24	.30	.00	.27	.07	— .19	.13
6. Role enactment	.77	.10	— .27	— .20	— .15	— .43	— .03	— .11	— .03	.25
7. Production	.67	— .09	.12	.06	.67	.09	— .16	.15	.14	— .03
8. Consideration	.39	.76	— .17	— .10	.08	.06	— .14	— .38	— .15	— .20
9. Predictive acc.	.42	.46	.74	— .20	.00	.01	— .16	.02	— .01	.01
10. Reconciliation	.51	.60	— .29	— .28	— .15	.00	.00	.42	— .12	— .08
% Variance	33.8	27.0	8.1	6.2	6.7	5.5	4.0	4.0	2.8	1.9

loading of .27 on Tolerance of Freedom. The factor is essentially a measure of *Initiating Structure*.

Factor VIII is found to have weightings of —.38 on Consideration and .42 on Reconciliation of Conflicting Demands. The minister who is perceived as high in demand reconciliation is perceived as low in Consideration of the members of the congregation. This result appears to be in accord with findings from research on personality and perception. Perceptual rigidity tends to be associated with high efforts to exercise control over relevant aspects of the environment. Factor VIII is identified as *Control*.

Factor IX is a specific factor, with a loading of .40 on *Tolerance of Uncertainty*.

All subscales have negligible loadings on Factor X. The highest loadings are .25 on Role Enactment and —.20 on Consideration. The factor does not appear to merit a distinctive title.

Four of the subscales (Representation, Tolerance of Uncertainty, Structure, and Predictive Accuracy) appear with at least moderately high loadings on specific factors. The most promising of these, for further development as an independent subscale, is Predictive Accuracy.

#### F. DISCUSSION

The use of 1.00 rather communalities in the diagonals creates certain difficulties in determining whether a specific factor really measures any specific variance. If it is assumed that the square of the loading of a subscale on a factor must assume a value larger than the difference between one (1.00) and the computed reliability of the subscale, then only Representation in Factor III, Tolerance of Freedom in Factor IV, and Consideration in Factor V meet this criterion for community leaders. Following the same procedure for ministers, only Predictive Accuracy in Factor III meets the criterion. Between the two populations, four of the subscales appear to account for some specific factor variance. Judged on this basis, the argument is not strong for retaining the separate identities of the subscales.

But there is another consideration that favors the retention of the separate subscales. The composition of the general factor differs considerably between the two populations. Only Persuasiveness, Initiating Structure, Role Enactment and Production Emphasis appear with higher loadings on the general factor than on other factors for both ministers and community leaders. Only three of the subscales (Representation, Initiating Structure, and Predictive Accuracy) appear on specific factors in both populations. Persuasiveness is associated with low Role Enactment among both community leaders and



ministers. Demand Reconciliation appears with low Role Enactment for community leaders, but with low Consideration for ministers. Except for the three specific factors that are common to both populations, the composition of the factors differs for ministers and community leaders. For this reason, it would be difficult to develop factor scoring keys that would apply equally well to both populations. The findings suggest that if the Leader Behavior Descriptions are to be used for comparative studies across populations, there is merit in retaining the identity of the separate subscales, and in attempting to strengthen the identity of each. Used in this manner, the new scales offer some hope of providing interesting, and perhaps useful, insight into the structure of leader behavior.

### G. SUMMARY

Eight subscales, in addition to Consideration and Initiating Structure, were developed for the Leader Behavior Description Questionnaire. The questionnaire was used to obtain descriptions of leaders in community development activities ( $N = 57$  persons) and ministers of various denominations ( $N = 103$  descriptions).

Factor analysis of the intercorrelations among the ten subscales produced the following findings.

1. Only four of the subscales (Persuasiveness, Initiating Structure, Role Enactment, and Production Emphasis) were highly weighted on the general factor for both populations.
2. Three of the subscales (Representation, Predictive Accuracy, and Initiating Structure) appeared with at least moderately high weightings on specific factors in both populations.
3. Demand Reconciliation appeared in the same factor with low Role Enactment for community leaders, but with low Consideration for ministers.
4. High Persuasiveness was associated with low Role Enactment for both community leaders and ministers.

### REFERENCES

1. BASS, B. M. *Leadership, Psychology, and Organizational Behavior*. New York: Harper, 1960.
2. BORGATTA, E. F., COTTRELL, L. S., & MEYER, H. J. On the dimensions of group behavior. *Sociometry*, 1953, 19, 223-240.
3. CARTER, L. F. Evaluating the performance of individuals as numbers of small groups. *Personnel Psychol.*, 1954, 7, 477-484.
4. FLEISHMAN, E. A. A leader behavior description for industry. In R. M. Stogdill, & A. E. Coons (Eds.), *Leader Behavior: Its Description and Measurement*. Columbus: Ohio State Univ., Bur. Business Res., Monogr. No. 88, 1957.

5. HALPIN, A. W., & WINER, B. J. A factorial study of the Leader Behavior Descriptions. In R. M. Stogdill, & A. E. Coons (Eds.), *Leader Behavior: Its Description and Measurement*. Columbus: Ohio State Univ., Bur. Business Res., Monogr. No. 88, 1957.
6. HEMPHILL, J. K. Situational Factors in Leadership. Columbus: Ohio State Univ., Bur. Educ. Res., Monogr. No. 32, 1949.
7. HEMPHILL, J. K., & COONS, A. E. Development of the Leader Behavior Description Questionnaire. In R. M. Stogdill, & A. E. Coons (Eds.), *Leader Behavior: Its Description and Measurement*. Columbus: Ohio State Univ., Bur. Business Res., Monogr. No. 88, 1957.
8. SHARTLE, C. L. Introduction. In R. M. Stogdill, & A. E. Coons (Eds.), *Leader Behavior: Its Description and Measurement*. Columbus: Ohio State Univ., Bur. Business Res., Monogr. No. 88, 1957.
9. STOGDILL, R. M. Individual Behavior and Group Achievement. New York: Oxford Univ. Press, 1959.

*Bureau of Business Research*  
*Ohio State University*  
*Columbus 10, Ohio*



## FORMATION AND EVALUATION OF HYPOTHESES IN DREAM INTERPRETATION\*

*University of Chicago and Chicago Institute for Psychoanalysis*

---

ERIKA FROMM AND THOMAS M. FRENCH

---

In an earlier paper Fromm (1) discussed French's (2) approach to the interpretation of dreams. French attempts to reconstruct and visualize the functioning of the Ego at the moment of dreaming through study of the dream work. This complements the classical approach of Freud. Freud's interest was focused on the dream as a distorted and disguised fulfillment of an infantile wish. French, starting with the classical approach, supplements it by studying the dream as a cognitive, integrative attempt of the Ego to solve a *current, Here-and-Now* conflict. French has called this the "Focal Conflict." This focal conflict is always based ultimately on one or more infantile conflicts, although in a particular case it may not be easy to trace it back to its roots in the dreamer's past.

### I

The authors consider dreams to be artistic products of the thinking, groping mind. The dreamer uses simultaneously ideation and more formally organized logical thought in producing the dream, just like the poet or the composer, or the painter do in their creative activities. And dreams, when once fully understood, are fascinating works of beauty.

In order to understand a dream, the interpreter must involve himself open-mindedly in a parallel re-creative, non-schematic process. It is an intuitive activity. But if it is to be more than a hit-or-miss spouting off of the analyst's own brainstorm it requires also scientific self-discipline and the willingness to evaluate critically and conscientiously the ideas and hypotheses one has arrived at intuitively. Only then can the interpreter find out whether his intuitive hypotheses are correct, need modification and/or refinement, or whether they should be discarded because they are wrong.

At the start of a dream interpretation the interpreter is in a situation similar to that of M. Broussard when he came upon the Rosetta Stone. The message of the dream—or parts of it—are written in hieroglyphs (the

---

\* Received in the Editorial Office on June 6, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

language of the patient's Unconscious) on a slab of black basalt. It must be translated faithfully into consciously understandable language (the language of the Conscious). How does the interpreter go about this task?

The interpreter should not just decode the "symbols." Like the *good* translator of poetry from a foreign language and culture, who tries faithfully and artistically to recreate in the language of the translation the specific poetic atmosphere and quality of the original poem, so the dream interpreter must also recreate the dream's specific, elusive atmosphere in order to understand it fully and make it meaningful to the patient's conscious mind.

Sometimes the interpreter intuitively grasps the central meaning of a dream. He then must proceed to see whether all the details fit in as parts of the whole—with his hypothesis. (If they do not, he must discard his hypothesis as being wrong, or modify it.) The following dream and its interpretation may serve as an example of an intuitive (re-) creative, correct hypothesis:

### A. THE CEMENT DREAM

#### 1. *Dream*

In the fifteenth analytic hour a 35-year-old male married patient,<sup>1</sup> father of three children, who came to analysis in order to be cured of his neurodermatitis, reported that a few nights ago he dreamt he "was mixing cement." That was all.

#### 2. *Associations*

Immediately upon awakening he felt as if he really knew exactly how to mix cement. But actually, he says, he does not.

He associates in the analytic hour that his house recently needed tuck-pointing; that he was going to do it himself, but did not know how to mix concrete. So, he hired somebody. The patient further associated: "My dear wife went to a bingo game last night." He had to take care of the children, put them to bed, and wait till she got home. Because she went out to play bingo, he came too late to his night job.

Then he talked about a hit-and-run driver who killed a child. The patient "cannot see how people can leave the scene—I couldn't do such a thing."

Later he tells of his wife's accusing him of wanting to get rid of everything, not enjoying his family, just tolerating them.

#### 3. *Interpretative Process*

That the hiring somebody might refer to the analyst whom he had just hired recently to help him with his problems, was easy to figure out (on

<sup>1</sup> This patient was analyzed by Dr. David Hamburg. The authors express thanks to him for allowing them to use the notes he took during the analytic hours for the interpretations they are making in this paper.



the basis of parallels). And the dream in connection with the first association upon awakening stated clearly that "mixing cement" was a "problem" that the patient felt he could not solve alone (again a parallel).

But *what* is the specific problem that the patient wants the analyst to help him with at this time? That is the basic question we had to answer in interpreting this dream.

In order to do so we let our imaginative processes play over the material. We visualized a house, tuckpointing, cement.

Suddenly it occurred to us that in "tuckpointing" cement is used to *hold the bricks of a house together*. The *function* of tuckpointing is to prevent the house from crumbling apart. . . . Our next thoughts were: A house is a home, a home for a *family*.

In a flash, we felt that now intuitively we may have grasped the central meaning of the dream; and we hypothesized that the patient's problem may be one of holding his family together.

This hypothesis immediately proved to be the key for understanding all the associations as well as the dream:

In all of the associations the patient refers to a wish to be a good family man . . . and a fear that he is not. He asks himself: Am I or am I not a good family man? He wants to be. He says his wife is the one who neglects her duties to the family and acts like a "hit-and-run" driver, he "couldn't do such a thing." But his wife accuses him of "wanting to get rid of" the family, just tolerating them. The dream expresses the patient's hope that in therapy the analyst will help him to hold his family together. The patient dreamt the dream in an *effort to solve* the problem of how to be a good family man. But this problem is expressed in the dream in a de-animated, de-vitalized symbolic form: instead of interacting with his wife and children directly, he "mixes (lifeless) cement."

## II

Rarely are a dream and its associations so concise, short, clear and simple that the dream's total latent meaning can be grasped by the interpreter at once, as a whole with all its constituent subparts.

Usually dreams are too complicated to be encompassed by the interpreter's integrative Ego span right at the beginning of the interpretive process.

This process starts in the majority of the cases with *scanning* of the material (usually followed by letting it "simmer" in the analyst's mind). Like a radar instrument, the interpreter lets his thoughts and feeling processes

sweep, in a random fashion, over the patient's dream material, the associations and the preceding hour, until he gets one or more clues.

Listening for clues . . . what are these clues?

There may be parallels in the material of the dream and the associations. Or incongruencies. Or a feeling of empathic resonance within the analyst. Or the interpreter may look first for the general affect of the dream which so often indicates the reactive motive. Or perhaps the totality of the situation in which the patient puts himself in the dream—or with which dream and associations seem to be pervaded—become a first clue to the interpreter.

Perhaps, on the other hand, the interpreter may be struck by strange, curious parts of the dream and get the feeling: this detail must mean something . . . but what?

### III

The *nature* of the clues often determines which particular method of further approach to the understanding of the dream the interpreter then follows. If he thinks he has caught on to the focal conflict as a whole, he must proceed to find the evidence in the constituent parts of the dream . . . evidence *for* as well as *against* his hypothesis. In Rorschach terminology, we shall call this approach from the whole to the parts the  $W \rightarrow D \rightarrow d$  approach.

Let us show on an example of modern art what we mean. Look at this picture by Joan Miro (see Figure 1, p. 275).

The first impression is one of lightness, grace, freedom, wit—on account of the mostly upward slithering, quicksilver, mischievous movement in the picture. The brilliant reds, blues, greens, yellows and whites in the actual painting contribute to the feeling of playfulness and gaiety. These are *W* impressions.

Then the viewer may become aware of details:

Clown-like cats which play with a ball of wool and pay no attention to a fish behind them. Tentatively the viewer makes an hypothesis that this could be a circus picture. He looks further and sees a ladder near which two people are swinging from a high trapeze. This confirms the circus idea.

Then he sees a jack-in-the-box who throws a ball into the left hand of a long-necked, round-headed man with a fancy mustache. The man holds a pipe in his mouth. His stomach seems to be a bouncing ball in a body contorted out of shape. What could that mean?

There are musical notes, and eyes in the painting. Why? the viewer asks himself. Could they symbolize that the act of looking and music are important features of what the artist wants to express? It fits with the circus idea.



The viewer then recognizes other animated, circus-like creatures that move and twirl about in amiable confusion. But he is still vaguely disturbed by the representation of the man whose throat seems to be so painfully elongated and whose stomach looks like a bouncing ball. He cannot make this jibe with the gay circus idea until he learns about one historical fact: Miro painted this picture when he was poverty-stricken and starving (4). Miro called the painting "Harlequins' Carnival."

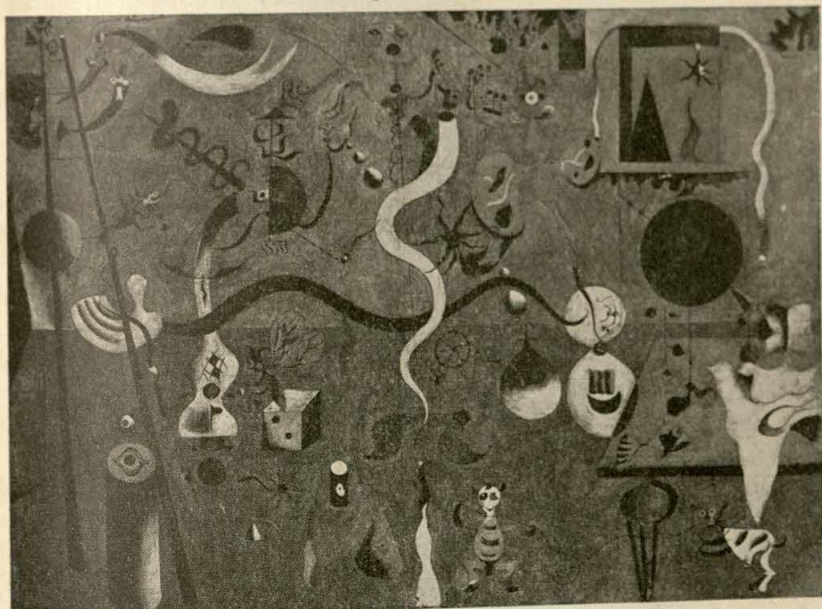


FIGURE 1

JOAN MIRO, HARLEQUINS' CARNIVAL (1924-25)

Room of Contemporary Art Collection, Albright-Knox Art Gallery. Buffalo, New York. Photo, courtesy, Museum of Modern Art, New York.

The danger in the  $W \rightarrow D \rightarrow d$  type approaches lies in the temptation to use a Procrustean Bed technique to make the details fit the hypothesis; or to discard as "unimportant" those details that do not fit in with one's hypothesis.

The greater the emotional commitment the interpreter has to his hypothesis the easier can he succumb to the Procrustean temptation. The scientist does get committed to his hypothesis. It's his baby. He has great Ego involvement in it (5). Anyone who has ever gone through the emotional process of

hypothesis making—with its hours of “labor,” toil, happy excitement, despair, and utter joy—can have no doubt about this.

If the initial clue has been but a detail—a minor (d, Dd) or major (D)

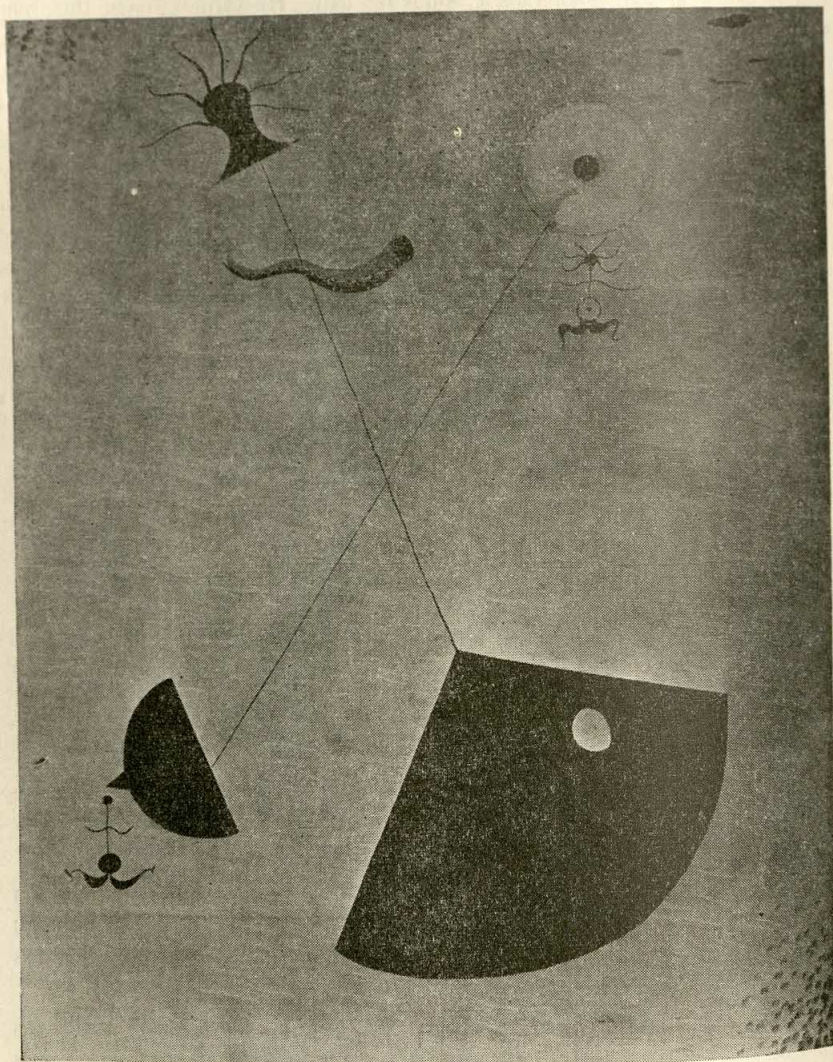


FIGURE 2

JOAN MIRO, MATERNITY (1925)

Collection Mrs. Roland Penrose

Photo, courtesy, Museum of Modern Art, New York.



piece of the mosaic of the dream pattern—the interpreter will look for other (meaningful) pieces until he recognizes the total Gestalt of the dream; usually with a joyful “Aha Erlebnis.” In these cases he proceeds from d or D to W.

Let us—as a simile—show you another picture by Miro (Figure 2). You cannot begin to understand the idea that Miro wants to express in this picture until you recognize that the lower left hand detail is a female breast in profile—with a small human figure close to it. Then, perhaps, you go to the upper right hand design and figure out that it could be a breast seen full face—and a child attached to it. It is an infant. The painter has given it typically infantile random movements. The lower right and upper left designs you now comprehend as being symbolic representations of breast and child, too, and the fine-spun lines connecting the elements in the opposite corners as indicating a *relationship*. After understanding all of these details, you now can “interpret” the whole, you say to yourself: oh, Miro wants to depict the idea of Maternity (and this, indeed, is the name of the picture). You still have a little trouble with that golden horn. It’s a horn of plenty, symbolizing the rich experience that maternity is for mother and child. It fits in with and enriches the interpretation given.

The interpretative process here described in making a surrealist painting intelligible is the same as the one the dream interpreter goes through when he deciphers a dream by means of the  $D \rightarrow W$  approach. We shall use this approach in:

## B. THE ARMY MESS HALL DREAM

In the ninety-eighth hour, the analyst praised the patient whose Cement Dream we discussed above, for having been able to allow himself to become aware of tender feelings towards his wife. The patient expressed pleasure and embarrassment in relation to this compliment at the end of the ninety-ninth hour. The next night the patient had a dream.

### 1. Dream

I was in the Army again. Went into the mess hall with another guy. Had a little trouble getting food. Then saw a couple of guys I recognized. A Southerner, called “the Colonel”—he was really a private—said: “Sit down, I’ll take care of you. I’ll serve you like a king, on gold plates.” I replied: “Never mind the plates, just get me good food.” Then this guy’s appearance changed . . . he looked like Jack Benny. He said he was Jack Benny’s brother. I said: “Are you related?” Everybody laughed. Then the scene changed. We were both being chased. We were going through a tunnel. I said: “This is ridiculous! If you are Jack Benny’s brother and he owns millions, why doesn’t he help out?”

## 2. *Associations*

The patient associated that he had been close to "the Colonel" in the Army. The "Colonel" had been willing to do quite a bit for him.

The gold plate business seemed odd to the patient. He said he "Didn't care for a lot of bullshit, just facts—don't flower it up. Sometimes people will give you a medal and then lay an egg."

He returned to thinking about the "Colonel," and about a lieutenant in the Army (a medical officer) for whom the "Colonel," worked. This doctor was rough. He said the patient was goldbricking, using his dermatitis as an excuse for discharge. The patient denied it. The lieutenant's superior officer overruled the lieutenant and got the patient discharged.

## 3. *Interpretive Process*

The manifest content of this dream as a whole at first seems understandable. Perhaps the interpreter is struck by a curious detail of the dream, the "Colonel's" words: "I'll take care of you . . . will serve you on gold plates."

It is incongruous to be served on gold plates in an Army mess hall. The juxtaposition of these two details seemed odd to the patient, too. Obviously they must be an important clue to the latent content of the dream.

The interpreter may make a mental note that an Army mess hall is a stark, uncozy, purely masculine place; a place where men use rough language (see the word bullshit in the associations) in order to hide or disparage sentiments.

Then the interpreter's attention may turn to the patient's next association: "Sometimes people will give you a medal and then lay an egg." A curious association. It sticks out like a sore thumb. To what perceptions, thoughts or feelings could the patient refer by quoting this saying?

As a working hypothesis, the interpreter may begin now to ask himself: does the patient symbolize in the dream the thought that someone who serves him is trying to flatter him? (Serve him—a mere army sergeant—"on gold plates like a king"; "flower it up"; "give a medal.") Who could the flatterer who takes care of him be? The analyst? Very possibly so. The analyst, after all has praised the patient recently about being able to express tender feelings towards his wife.

Now the association "Sometimes people will give you a medal and then lay an egg" falls into place and a part of the dream's puzzle becomes intelligible. The patient is pleased by the analyst's praise, but also intensely embarrassed about his tender feelings; and he fears the analyst's praise will lead to something unpleasant. What this unpleasant consequence is, we don't



know yet. Nor can we be *fully* sure at this point that the "'Colonel' who really was a private" stands for the analyst.

Why does the "Colonel" change into the brother of Jack Benny, i.e., the brother of a famous and rich man? And why is he chased out through a tunnel together with the patient? We could find no answer to these two questions and thus could not understand the last half of the dream until it occurred to us that the young analyst was the Chief Psychiatric Resident at the clinic at which the patient was seen, that he occupied an office next to the office of the Chief Psychiatrist who was a famous man, and that a tunnel connected the Psychiatric Clinic with other parts of the hospital—and led out to the street.

We can now be certain that the "'Colonel' who really was a private," and whose physiognomy changes so he looks like Jack Benny's brother represents the analyst. What is more, now the total latent dream content forms an organic, intelligible, cognitive whole.

The patient felt pleasure at being warmly praised by the analyst for expressing gentle tenderness towards his wife. But he blusteringly must deny this in the dream because he also feels shame. According to his Teutonic ego ideal, a real he-man should not be tender. Therefore, in the second half of the dream his thoughts turn to depreciating his analyst's importance. ("He said he was Jack Benny's brother. I said: Are you related? Everybody laughed.") Furthermore, the patient fears that the hospital authorities will chase him out of treatment—and the "sissy" analyst with him—if they hear that this resident who takes care of him has "flattered" him for being able to express tenderness. As the patient's masculine pride disapproves of gentle tenderness within himself, so he thinks the men in charge of the hospital must do; and the director of the psychiatric clinic, his resident psychiatrist's big brother, "doesn't help out."<sup>2</sup>

#### IV

Which method of approach the interpreter will use in solving the riddle of a dream is an intuitive decision.

In some cases the nature of the dream material itself "demands" a particular interpretative approach. In trying to open a heavy, tightly nailed-down wooden case you don't use your finger nails, you use a crow bar—that's what "the thing asks for."

In other cases the manner of approach may be determined by subtle factors

---

<sup>2</sup> The transference meaning of this dream will be discussed in our forthcoming book, *Method of Dream Interpretation*.

residing within the interpreter's own personality structure—and we don't mean counter-transference. Neo-Gestalt theoretical research on perception (Klein (6), Witking et al. (7)) and on cognition (Klein (8), Scheerer (9)) has shown that, within certain subtle limits, personality characteristics and/or the particular psychic and somatic state a person is in (e.g., hunger) determine whether and how he perceives an objective stimulus or a problem. It is for these reasons that certain clues may immediately be perceived (in a certain context) by one interpreter, and not by another.

Also: on the basis of their own personality structure and their characteristic mode of working, some interpreters may naturally tend to use one interpretive approach rather than another.

Frequently, without necessarily being consciously aware of it, one uses several methods in succession . . . or in combination . . . in the interpretation of one dream.

There is only *one*<sup>3</sup> fully correct Focal Conflict interpretation for every dream . . . but there is not just one road that leads to it. Important is only that one gets there, and is not satisfied with being merely somewhere in the vicinity of the correct understanding.

## V

The question arises now: What makes for failure in dream interpretation?

Failure can be due to the interpreter's unwillingness to evaluate critically his intuitive hunches. It occurs when instead of checking and rechecking for correctness, he deifies his intuition and allows no doubt about it. If challenged, some analysts will burn at the stake those that advocate scientific controls.

In the more enlightened, less fanatic believer in "pure" intuition, failure to find the Focal Conflict frequently is due to his getting side-tracked on a Sub-Focal Conflict.

Another possible cause of failure may be the interpreter's commitment to an (incorrect) hypothesis he has made at an earlier stage of the interpretative process, and his resistance to letting go of it. This point was touched upon above (page 275) but deserves further consideration here:

An intuitively made hypothesis is the interpreter's artistic creation. He has a great deal of Ego involvement in it. The *purely* intuitive analyst makes *one* hypothesis. He puts all of his eggs in *one* basket. Naturally, he is not willing to drop it.

<sup>3</sup> It is true that there are many overdetermined meanings of a dream. But these meanings vary greatly in the intensity of their cathexis at a particular time. The Focal Conflict is the one that it most intensely cathected at the time of dreaming.



The intuitive interpreter with the Scientific Critical Attitude—he who looks for evidence that could prove or disprove his hypothesis and who also looks for possible gaps—is committed to his hypothesis too. But he is *mainly* committed to his hypothesis, *not solely*. Some of his Ego involvement is taken up by the sporty pursuit of proving his own hypothesis wrong or insufficient.

Actually, after his hypothesis has occurred to him he makes two more: (a) my hypothesis is correct; and (b) it may be wrong; with the major Ego involvement, of course, on (a).

Because he is not totally committed to his one artistic-intuitive creation, he has some free-floating "committability" left and is able to evaluate his own creation critically, to let intuitive emphases shift, and to find new solutions if necessary.

Still another cause for failure is the fact that so frequently an important clue to the dream lies right under the interpreter's nose but is not recognized as a clue. Or it is recognized only after a good deal of "simmering," incubating about the dream—and shifting of emphasis for and against certain hypotheses the interpreter has made. The clue has been "camouflaged" as a subpart in the context of another cognitive whole. Suddenly the overlooked clue jumps out of its camouflaged position and stands out by itself (or ties up with another part into a more meaningful whole). And the interpreter asks himself: Why, in Heaven's name, didn't I see that before?

Scheerer and Huling (9) published a beautiful paper on "Cognitive Embeddedness in Problem Solving" which deals with the same topic in near-to-ordinary-life problem solving situations.

## VI

How, in the end, can the interpreter be sure that he has faithfully, i.e., correctly, caught the Focal Conflict and the exact total meaning of a dream?

The answer to this question is: The interpretative hypothesis is correct if it illuminates or explains every bit of the dream text plus every single association given to it as parts of *a total response the dreamer has made to his actual current emotional situation*. Fidelity is measured by the degree to which the interpretation approximates the ideal of such total illumination and explanation.

Two types of checks are possible. One is of scientific nature. The other is an artistic kind of validation which we shall call the Esthetic Experience.

In the process of seeking scientific evidence, the interpreter carefully,

minutely, checks and rechecks the material for inconsistencies with his hypothesis. The interpretative hypothesis about the dream is correct, if in the end the interpreter does not need to resort to subsidiary or extraneous side hypotheses to tie up loose ends; and furthermore, if he has been totally honest in his critical evaluation as to whether or not all the material fits in with his hypothesis. He must not violate any of it by using a Procrustean Bed Technique.

The other evidence for the correctness of an interpretation is a more subjective artistic experience the interpreter has with some—not all—interpretations he makes. This is what we propose to call *the Esthetic Experience*.

The Esthetic Experience is a close relative to the Aha-Erlebnis, but not its identical twin.

What happens in the interpreter—who may be the therapist, or the patient, or preferably both searching together—while he is trying to solve the riddle of the dream?

He has found some clues. Perhaps has brought some or even most of them into a meaningful relationship with each other. But there has remained some part of the material that has defied explanation. Or the whole did not make “real sense” to him. The analysts sits back and “simmers,” lets his radar-like, freely suspended attention wander over the dream again.

Suddenly he arrives at a new or modified artistic insight about the dream. And feels: Oh, . . . that’s it! That’s beautiful! That illuminates the whole meaning of the dream and I now can see it in all its plasticity as the beautiful work of art that it is. He feels exhilarated, excited, joyous, and experiences esthetic pleasure.

Is the Esthetic Experience the interpreter has, real proof that he has arrived at the correct interpretation? Subjectively it certainly feels like it. And frequently it really is. Particularly, if the patient comes to this sudden insight or if the therapist’s interpretation “hits home” with the patient and, like a sunburst, illuminates to him too, the dark landscape of his unconscious struggle.

When the interpretation is made in the process of research—from a written record, post hoc, in the absence of the patient—some additional, consensual proof can be gained if the esthetic feeling is experienced by your colleagues when you present to them the patient’s dream and your interpretation. We hope you felt it when we gave you the Cement Dream and our interpretation of it. And perhaps some of you felt only what we should like to call the *Hm-hm Erlebnis*. This can be described as the arriving at an insight which proves someone else’s hypothesis rather than one’s own. It contains insight



. . . nodding-agreement insight. But it is not as creative an insight as that experienced in the Aha-Erlebnis and in the Esthetic Experience.

Nearly always the Esthetic Experience is subjectively felt by the one who experiences it as pretty strong proof. But it *can* have its pitfalls: The investigator may get so enchanted by it that he fails to fincomb the material for details that might not fit in with his hypothesis. The Esthetic Experience does not necessarily suffice. It always must be followed, at least for check-up purposes, by the process of finding the Scientific Evidence.

Much of what has been said in relation to dream interpretation here perhaps applies to all creative scientific hypothesis formation and evaluation.

#### REFERENCES

1. FROMM, E. The Psychoanalytic Interpretation of Dreams and Projective Techniques. *Am. J. Orthopsychiat.*, 1958, **28**, 67-72.
2. FRENCH, T. M. The Integration of Behavior. Vol. I-III. Chicago: Univ. Chicago Press, 1952, 1953, 1958.
3. FREUD, S. The Interpretation of Dreams (1900). The Basic Writings of Freud, New York: Modern Library, 1938.
4. ERBEN, W. Joan Miro. München: Prestel Verlag, 1959, p. 119.
5. ROE, A. Creativity and the Scientist. Lecture given at the Ill. Psychol. Assoc. Conv. Chicago, October 9, 1959.
6. KLEIN, G. S. The Personal World through Perception. In Blake, R. R., & Ramsey, G. V. (Eds.), *Perception: An Approach to Personality*. New York: Ronald Press, 1951, pp. 328-355.
7. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. Personality through Perception. New York: Harper, 1954.
8. KLEIN, G. S. Cognitive Control and Motivation. In Lindzey, G. (Ed.), *Assessment of Human Motives*. New York: Rinehart, 1958, pp. 87-118.
9. SCHEERER, M., & HULING, M. D. Cognitive embeddedness in problem solving. A theoretical and experimental analysis. In Kaplan, B., & Wapner, S. (Eds.), *Perspectives in Psychological Theory*. New York: Int. Univ. Press, 1960, pp. 256-302.

*Department of Psychology*  
*University of Chicago*  
*Chicago 37, Illinois*

## PATIENT WISDOM: AN ANTHOLOGY OF CREATIVE INSIGHTS IN PSYCHOTHERAPY\*

*Western Psychological Center, Sherman Oaks, California*

---

STEWART B. SHAPIRO

---

### A. INTRODUCTION

There are times when I feel like putting down my clinical notebook and just listening to what patients have to say. This, you might say, is what all good therapists do all day long anyway. So, does this justify a paper?

What I mean specifically by listening is an attitude of openness to hear and to sense what is wise in what patients say. My major purpose in this paper is to explore the impact of patient wisdom on the therapist. To do this I find it necessary to temporarily set aside the clinical-scientific frame of reference of a psychologist, and to relinquish for a while my precious, hardwon skills of analyzing, categorizing and reconstructing.

### B. PEAK EXPERIENCES IN PSYCHOTHERAPY

I suppose that I could easily make a case for the idea that there are certain relatively rare moments of revelation and insight in psychotherapy, when the patient is really "with himself." And it could easily follow that these moments are the essence of psychotherapy, and that they almost always represent therapeutic movement. These are the peak experiences, the nodal points when both my patient and I feel something is really happening—something good.

It is difficult to drop the stance of psychologist, once having assumed the position professionally *or* as an amateur. As I listen I have little doubt that these moments of patient wisdom do, in fact, represent therapeutic movement, but let us lay aside, for now, the therapeutic significance of these bits of human expression, and respond to the sheer wisdom of patients as revealed in the therapy room.

### C. CRITERION OF WISDOM AND SELECTION OF QUOTES

First, before we read, a word about the criterion of wisdom and the selection of the quotes which follow.

---

\* Received in the Editorial Office on June 6, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



Twenty-one sample statements were judged as "wise" by six professional therapists, including myself, on a three-point rating scale. These statements were selected from all of the recorded productions of ten patients in long term individual psychotherapy with me. Most of the statements were reconstructed from notes. Others were taken verbatim from tape recordings.

#### D. SELECTIONS ABOUT THERAPY AND THERAPISTS

1. In therapy like in all other human relationships, it isn't a matter of who is sick and who is well—there's much more to it than that. It's much deeper. It's a driving need. I appeal from strength for strength. I appeal for help to peel away the 95 per cent intellectual jazz—the crap—and get down to something rich and simple—flowing and bubbling—like love.

2. If I'm sick and *you make me well*, then I'm no longer a "me" person. It's an ersatz sort of thing—everything is taken away from me and goes to you. In "succeeding" I have to be me, if I'm not, I don't get satisfaction. It's like another voice talking and I *have* to be *me* and get satisfaction by *not* succeeding.

3. I've got a new development in my thinking about therapy and therapists—a dethronement. I've always *felt* it was a religion even though I *knew* better. Suddenly I had the feeling it's a very limited thing. This is a very minor business. I don't have the feeling it's a world-shaking event. I can't depend on therapy for my life. You can't learn how to *live* from therapy.

4. Therapy is wallowing in shit—you have to examine it—pick it up with your hands—smell it and find out it's not as bad as you thought it was.

#### E. ON THE MEANING OF PSYCHOLOGICAL SYMPTOMS— SECONDARY GAINS

1. Something very interesting happened here. Ever since yesterday I've realized that feeling pretty good about myself is a threatened separation. It had a lot to do with the picture about what martyrdom means. It was one step beyond sadism. I'm impressed with the solution. I have respect for the kid that dreamed it up—the martyrdom I mean. I kept running into myself and realizing I was feeling *different* and very good—the closest thing to elation—it has a lot to do with independence—and then comes the possibility that I'll be abandoned because I'm not pathetic enough.

2. I'm trying to prove to somebody I'm sick and I use it so I don't have to stand up and say outspokenly what I feel and be responsible for what I do.

#### F. ON THE MEANING OF LIFE—LIFE IS A STRUGGLE

1. Ever since I've been "up" I've noticed I've accepted the fact that, Goddammit, life is a struggle—I'm willing to give what it requires

to achieve the goals I want—it's a struggle just to breathe sometimes—(laughs) but I feel a hell of a lot better when I'm struggling. I feel alive! Maybe that's my role—that Goddammit—it's not gonna be easy for me. I feel better when I'm struggling. I used to lie down and wonder why I was here, what's it all about? What difference does it make why I'm here, I'm here and I want to make the most of it. For the first time this struggle has some meaning for me.

2. I'm just beginning to see people are not like me. They are different as people—a new awareness of their struggle to "become"—to be recognized, to discover, to orient, to take hold—a feeling they really are terrible loose—floating around, bouncing around. They work terribly hard to figure things out. I'm aware of their tension, sensitivity, vulnerability and ruthlessness. I think of them as floating, trying to grasp the walls and maybe the floor a little bit—trying to get themselves a solid position and they need a lot of encouragement.

3. I hit on a new idea today—the idea that life doesn't have any meaning. There's no use fighting it—but people can bring meaning to it. It's not foolproof, and it isn't easy either. I was expecting this idea to do something for me. I worked it to death. It just doesn't work that way. You can't just find an idea, sit back and relax and not try anymore. The real trick is not waiting for an idea or for somebody else to give life meaning. What made it possible for me was that there *was* nobody to give it to me. It's still a matter of putting on the coals, one by one. You can build up a rich fire that would last all night.

4. It's not the fact that you are dying that's so bad, but it's the sudden terrible, frantic, realization that you've never lived.

5. I see old people. Who would dream old people are human? Old piles of rubbish—old, wrinkled and gray. Why stick around for that? Yet the dumb motor pushes on . . . it's a poor bargain. Life cuts you to pieces and every once in a while it gives you a lollipop. I feel cheated, like I'm the victim of a horrible hoax or joke and it all seems so senseless, but I'm in awe of other people . . . struggling, flinging back the hoax . . . why do they do it?

6. If you don't have your own ego filled up you try to rob someone else's. You get people down to your size. You destroy.

## G. INTERPERSONAL RELATIONS, INDIVIDUAL-SOCIETY, MAN-WOMAN, PARENT-CHILD

1. I hate this damn unyielding society—slaps you around, makes you a misfit and then punishes you for it—double jeopardy. It's just so grossly unfair. I don't understand why people like me grow up to be law-abiding citizens—no guts, that's all—fear of more punishment. I have a feeling that society owes me a period of lawlessness—run over the foster mother—let them be completely obnoxious—terrific demands—no "please," "thank you," etc. To be polite at that point for somebody like me is a further insult. I ought to be able to act like a baby. After



a period of that I'd be willing to melt in. I've got some vengeance to wreak. I've got a lot of "to hell with you(s)," "kicks on the shins," and "I don't give a damns" to get off my chest.

2. A good woman has the ability to give a man the experience of being like a mother—or being like a woman. In order to treat a woman like she should be treated, he has to be like a woman and to accept being like a woman. This is the one thing a man can't have, and yet he can. That's the one thing a woman can give up to him.

3. The way not to be outgrown is to go to your child and learn from this child.

4. When I really make contact with another human being, my problems just *start*. It doesn't *solve* all my problems like I always felt it would.

5. I'm just not "me" with him. It's so hard to admit I don't have the strength. I despise myself for not being like him. I can't accept him as a man and myself as a woman, it scares me—and yet to accept him as something different from me makes us both the same—it makes us close.

#### H. SELECTIONS ON THE INNER SELF

1. Fire is uncontrolled. It doesn't want to control itself. It leaps up and is comfortable—colorful. It would be nice to be a fire—lovely and free, shapeless and uncontrolled. What gets burned doesn't matter. Water humiliates a fire—this magnificent force—so sure of itself—so unaware that anything can destroy it. It's crushing. It ruins everything but water is cold and unfeeling.

2. I feel all revved up—like I'm sitting listening to my own motor. I'm aware of my own power—an increased awareness of my intelligence, potential and curiosity about things in general—and yet I have a fear of sinking—being sucked away to oblivion—an undefended me—raw—all the hard shell gone. I'd like everybody or at least somebody to know what a hungry defenseless person I am—that I *need* love. I don't want to be hard and pretending that I don't need. I dislike the hard person. It's not easily destroyed, but it's not getting anywhere either. I'm tender and open and liable to grow like tender tissue. The shell has become conscious and turned its eyes inward to see what's inside—to see what it has defended all this time. The shell is also hungry for growth, but not so damn sure it's OK to leave "the rest" of me.

3. I hark back to an old time when I was right. I was unjustly treated as a child. I feel right. The victory is owing that child, *not* the adult now. Make this child right. I want to feed the courageous child. I'm very proud of that particular "me" that wants to win victory for the child—it's sweet, covered over and not aware of being in therapy. It goes on living without being aware that it lives.

4. The little girl in me wants to grow up and become me—or get close to me through Tom. It's like finding yourself through somebody else. Once we admit we want to be one another, it's like a confession, then I can go and reach for him. I need contact to be intact.

## I. GENERALIZING QUESTIONS AND HYPOTHESES ON PATIENT WISDOM

After a period of reflection and appreciation of the human wisdom of these statements, I find myself reacting to them with three (psychologist) questions. First, what *is* wisdom? Second, what do all these expressions have in common? Third, what does this teach us about therapeutic change?

### 1. *What is Wisdom?*

This question is in line with Maslow's (2) recent work on qualities like wisdom, truth, goodness, honesty, etc. Interest in these human qualities has been generated through his work on self-actualizing individuals and peak experiences.

Webster's Unabridged Dictionary suggests that wisdom is sound knowledge, judgment and the *use* of the facts and experience of human life and conduct. This knowledge, judgment and action involves a system of values sensitive to truth and genuineness, and qualities of mental soundness and awareness.

In psychological language wisdom involves a value system, certain non-defensive intellectual ego functions, and certain characteristics of mental health. Wisdom, however, is more than intellectual functioning, more than a value system, and more than being mentally healthy. It involves a wholeness, or creative integration of all of these factors.

### 2. *What Do These Expressions Have In Common?*

Although the twenty-one statements in this paper range from broad philosophical comments and various value statements through sociological questions to psychological issues, they all seem to have the following characteristics in greater or lesser degree.

a. They involve a creative use of language and concepts, if not literary, then at least different from the ordinary syntax.

b. They are highly personal statements, not super-objective and distant. They are invariably in the first person, in tone, if not in grammar.

c. They are "emotionally loaded" statements. There is much affect, color, passion and depth of feeling, often but not always stated with great certainty or conviction.

d. They all involve a conflict and some kind of at least tentative creative approach to this conflict, not necessarily a solution. This is very similar to Barron's (1) finding on reconciliation of opposites and the search for novel solutions in highly creative professional people.

e. They have a certain kind of universality. They touch on problems of human existence, and might strike an empathic chord in many people.



### 3. *What Is Essential In Therapeutic Change?*

What do these kinds of statements teach us about the essence of psychotherapy? I'm struck by the possibility that in a sense, the patient may cure the therapist. Wordsworth said "the child is father of the man." My patient says "The way not to be outgrown is to go to your child and learn from this child."

With each patient a therapist learns all over again to identify his struggling child with that of the patient. The therapist learns again how to perceive growth, how to listen and understand. In a therapeutic atmosphere or a sustaining human relationship the patient unfolds his inner self and *corrects* the therapist's superficial socially conditioned perceptions and prejudices. In addition to this education in perception, the patient has a direct impact on therapist wisdom with his own wisdom. A therapist, through repeated contact with the sometimes painfully unfolding wisdom, may grow humble and wiser himself.

To paraphrase Kahlil Gibran on children:

Your patients are not your patients. They are the sons and daughters of life's longing for itself. They come through you, but not from you, and though they are with you, they do not belong to you. You may give them your love, but not your thoughts, for they have their own thoughts. You may house their bodies, but not their souls. You may strive to be like them, but seek not to make them like you.

### J. SUMMARY

The major purpose of this paper is to explore the impact of "patient wisdom" on the therapist in psychotherapy. To do this it is necessary to listen openly to the inner struggles of people, without analyzing, categorizing or reconstructing.

Twenty-one "wise" statements were selected and quoted from 10 patients of the writer. Three generalizing questions were derived from this procedure. These questions concerned the meaning of wisdom, commonalities in the statements and the essentials of therapeutic change.

It is suggested that wisdom involves a creative integration of values, non-defensive ego functions and mental soundness. These expressions were highly emotional and personal and yet retained a certain kind of universality. The writer concludes that one essential to therapeutic change is the therapist's receptivity to "patient wisdom."

## REFERENCES

1. BARRON, F. Implications of a Continuing Research Program on Creative Originality. Symposium, Western Psychological Association Meetings, San Francisco, 1962.
2. MASLOW, A. H. *Toward a Psychology of Being*. Princeton, N.J.: D. Van Nostrand Company, Inc., 1962. Pp. 214.

*Western Psychological Center*  
*4911 Van Nuys Boulevard*  
*Sherman Oaks, California*



## PERCEPTUAL FUNCTIONING IN THE DUCKLING\*<sup>1</sup>

*Department of Psychology, Queens College*

---

NICHOLAS PASTORE

---

### A. PURPOSE

The present experiments are a follow-up of Pastore's exploratory work (1958, 1959) on the perceptual functioning of ducklings. Further investigation was desirable because light-experience and interproblem transfer effects were not controlled. Moreover, sample size was small and the procedure inefficient.

Two considerations guided the plan of the experiments in securing evidence on the relative influences of empiristic and autochthonous factors on perceptual processes: the use of relatively mature animals at the time of hatching, and the onset of training at the earliest feasible age.

### B. METHOD

#### 1. *Subjects*

Sixteen incubator-reared Peking ducklings participated in the various experiments in groups of four. Although the inside of the incubator itself was dark, the actual hatching took place in a small dark bin. They remained in the bin until removed for experimental purposes. Moreover, they were returned to the dark bin in between experimental runs and after terminal feeding at the end of the day. If an *S* had not achieved criterion by the end of the experimental day, it was food-deprived in the bin until its turn came for further training trials on the next day. An experimental run consisted of a predetermined block of 20 trials. The number of daily runs was not fixed. The average reward, a mash of bread and milk, weighed about 0.1 gm per trial. If training extended into the third posthatch day, the reward was increased up to 0.3 gm at most. Training was begun in the first posthatch day for 14 *Ss*, and in the second posthatch day for two *Ss*.

---

\* Received in the Editorial Office on June 12, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> A grant from the Frank M. Chapman Memorial Fund, under the auspices of The American Museum of Natural History, supported the present research. I wish to thank Miss Marion Frank for assistance in running the animals and for editing a preliminary draft.

## 2. Apparatus and Procedure

A stimulus platform (seven inches long, two inches high,  $1\frac{3}{4}$  inches deep) was used in all experiments. Five shallow food wells were set into the top of the platform. The two end wells were angled slightly so that objects placed on them might have increased "attention" value. Previous work suggested that Ss tended to avoid the extreme objects. Positive photostats of figures or striations were glued on wood blocks for the stimulus objects of Problems 1 and 2 (Table 1), and black or white photostat paper for the stimulus objects of Problem 3. A block had a face one inch wide by  $1\frac{5}{16}$  inches high; its depth was  $\frac{3}{4}$  of an inch. Since the paper surface was protected with clear lacquer, occasional smudge marks left by an S's strike could be wiped away readily. Fresh stimulus objects were available for substitution for all problems.

*a. Problem 1 (triangle vs. other geometric figures).* Blocks with the outline figures in the same orientations as depicted in Figure 1 were prepared. Exclusive of the trapezoid, all figures were of equal areas. The set of figures

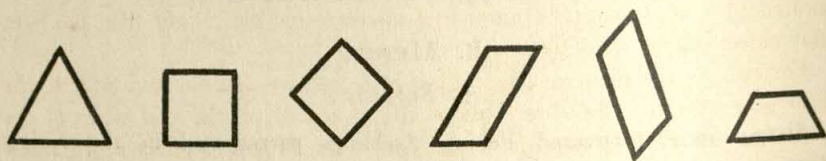


FIGURE 1

PHOTOSTATS OF ORIGINAL OUTLINE FIGURES

The length of a side of the triangle,  $\frac{3}{4}$  in., provides the scale for all other measurements.

was such that an isolated differential cue could not be the basis for a positive response to the triangle. For example, the acute angles of the parallelograms, trapezoid, and triangle were equal.

Five figures, exclusive of the trapezoid, were the discriminands for Problem 1. After criterion was achieved, there were two successive blocks of 20 trials each. In the first block, the trapezoid was substituted for the square. The aim was to determine whether a discrimination between the triangle and trapezoid had been effected. In the second block of trials, reversal training was undertaken. The square replaced the trapezoid and was made the positive figure. If the Ss responded to the square on a chance basis then odor would be excluded as an effective cue.

*b. Problem 2 (vertical vs. horizontal striations).* There were four blocks with vertical striations, four with horizontal striations. The striations were



black on a white ground. The width of each striation was  $1/32$  of an inch. Every block had four striations. For two Ss one block with vertical striations was contrasted to four blocks with horizontal striations; a block with horizontal striations was the positive figure for two other Ss.

c. *Problem 3 (black vs. white)*. Four black and four white objects were available. For two Ss one black object was contrasted to four white ones; a white object was positive for two other Ss.

d. *Problem 4 (brightness constancy)*. There were two blocks, one with No. 2 gray paper (light) and the other, No. 4 gray paper (medium). These numbers pertain to the gray series manufactured by the Color Aid Co. (New York). Numbers 2 and 4 respectively match Numbers 7 and 5 of the Munsell gray series. For the constancy trials, an opaque cardboard partition separated the two gray objects. The depth of the partition forced the Ss to make a choice between the objects at a distance of at least nine inches from the platform. A gooseneck lamp, with either a 60 W. bulb or a 100 W. bulb, was available for directing extra illumination to one side of the platform. Physical measurements and the independent judgments of the two Es with a reduction screen showed that, for either bulb, the medium gray object on the illuminated side reflected more light than the light gray object on the shadowed side.

For two Ss the medium gray object was positive and for two others, the light gray object. The objects were always placed on the end wells of the platform. In the first step of the procedure, the Ss had to effect a brightness discrimination. In the training of this discrimination, the distance of the platform from the cage was increased from zero up to the final distance of 31 inches. In the second step, the partition was placed on the platform and the Ss were given further trials in order to habituate them to the presence of the partition and the gooseneck lamp. After criterion was restored, the Ss were ready for the constancy trials. In the first block of 20 trials, the light beam from a 60 W. bulb was directed to one side of the platform and in the next block of 20 trials, the light beam from a 100 W. bulb.

e. *General points*. The Ss had to learn two initial tasks: (a) to orient themselves to the platform for food; and (b) to knock over stimulus objects for food. These tasks were quickly learned, and the details are omitted. During the course of training for Task b, the distance of the platform from the cage was increased from zero to 12 inches. The 12 inch distance was constant for all problems, except for Problem 4. Moreover, at some point in the training for Task b, the cage exit was blocked off in between trials so that an S was not visually stimulated by the operations entailed in rearranging stimulus objects on the platform.

It should be emphasized that the number of trials required for establishing proficiency in Task *b* were included in the count to criterion because all objects relevant to a given problem were on the platform. For example, with reference to Problem 1, one *S* reached criterion after 20 trials. These 20 trials were the very trials involved in teaching the *S* to knock over a stimulus object.

The principle of randomization applied to the placement of objects on the platform in all trials except for those trials in which some position response was to be broken. The learning criterion was 10 hits in a predesignated block of 20 trials for Problem 1 (five different objects were involved). The criterion was so chosen because then,  $|z| > 3$ . For all other problems the criterion was 15 hits in 20 trials (two different kinds of objects were involved). A correction procedure was followed throughout.

### C. RESULTS

The main results are set forth in Table 1. When the trapezoid was sub-

TABLE 1  
RESULTS

Problem	N	Trials to criterion		<sup>a</sup> Age criterion was reached for all <i>Ss</i>
		Mean	Range	
1. Outline triangle vs. outline geometric figures	4	55	20-120	1, 1, 1, 2
2. Vertical vs. horizontal striations	4	55	20-80	1, 1, 2, 2
3. Black vs. white	4	20	20-20	15 hrs., 1, 1, 1
4. Brightness constancy				
a. brightness discrimination	4	65	40-100	
b. constancy trials	4	5	0-20	15 hrs., 2, 2, 2

<sup>a</sup> Age is given in days, unless otherwise indicated. Age of one, for example, means that criterion was reached in the second posthatch day.

stituted for the square for the four *Ss* of Problem 1, three *Ss* continued to select the triangle at criterion level in the first 20 trials. Therefore, the response of the three *Ss* to the triangle was not determined by its two base angles alone. When reversal training was undertaken in the next block of 20 trials, the four *Ss* chose the square on a chance basis. Since in these trials the displacement of the square led to food, *Ss* did not respond to an odor cue in their original choice of the triangle.

Three of four *Ss* showed direct transfer from the initial brightness dis-



crimination to the first 20 trials of the constancy situation (60 W. bulb). In the next block of 20 trials all *Ss* performed at criterion level (100 W. bulb).

#### D. DISCUSSION

*Ss* had to learn several tasks in achieving criterion on a problem: adaptation to the experimental situation, orientation to the platform for food, knocking over a stimulus object, and the association of a particular stimulus object with food. Such considerations, in conjunction with the early manifestation of perceptual achievements, minimize the relevance of an empiristic approach and strengthen the cogency of autochthonous determination. Two specific results may be cited which pose a problem for an empiristic approach.

One *S*, at the age of one day, could discriminate a triangle from other closely allied geometric figures after the 20 training trials involved in teaching it to knock over an object. This *S* had spent one hour in the experimental situation. Certainly one may assume that "perceptual learning" occurred in the 20 training trials and in the one hour the *S* was exposed to the visual environment. Such an assumption may be construed as essentially equivalent to autochthonous determination, given the ambiguity of currently formulated empiristic theories in their application to the specifics of a perceptual problem (Pastore, 1956 and 1960).

Three of four *Ss* transferred the initial brightness discrimination to the constancy situation, a situation characterized by variable stimulus conditions in the presentation of objects. Under the conditions of the experiment, these *Ss* did not have the opportunity of learning to "correct their response" in the one-sided lighting which prevailed in the constancy trials.

Generally, the over-all results may restrict an empiristic theory to those animals born or hatched in relative immaturity.

#### E. SUMMARY

The perceptual functioning of 16 ducklings was investigated shortly after hatching. Each *S* was light-deprived prior to its entry to an experimental situation and between blocks of trials during the course of training. All *Ss* were assigned only one perceptual task.

Different groups of four *Ss* discriminated an outline triangle from other outline geometric figures, horizontal from vertical striations, and black from white objects. Brightness constancy was demonstrated in one group of four *Ss*. Criterion level was reached by all *Ss* at the median age of one day.

The relevance of the results to empiristic and nativistic approaches to perception was briefly discussed.

## REFERENCES

1. PASTORE, N. An examination of one aspect of the thesis that perceiving is learned. *Psychol. Rev.*, 1956, **63**, 309-316.
2. ———. Form perception and size constancy in the duckling. *J. Psychol.*, 1958, **45**, 259-261.
3. ———. Perceptual functioning in the duckling. *J. Genet. Psychol.*, 1959, **95**, 157-169.
4. ———. Perceiving as innately determined. *J. Genet. Psychol.*, 1960, **96**, 93-99.

56 Seventh Avenue

New York 11, New York



## REACTION-TIME AS A FUNCTION OF STIMULUS INTENSITY AMONG NORMAL AND PSYCHOTIC SUBJECTS\*

*University of Pittsburgh School of Medicine and  
Western Psychiatric Institute and Clinic*

---

H. E. KING

---

### A. INTRODUCTION

The existence of a systematic relation between states of human behavior disorder and reduced or otherwise faulty psychomotor functioning is, by now, well-established. The long history of investigation of problems of the kind (4, 5, 6, 7, 11, 13, 16, 17, 18) and the results of recent experimental approaches (1, 5, 7, 8, 10, 15) can leave little doubt that psychomotor adequacy is affected to a measurable extent whenever severe disruptions of the patterns of general behavior are noted, whether these be of physiologic, psychiatric or neurologic origin. The meaning of the phenomena is less clear. Speculation and logical or correlative analysis have been offered, proceeding often from the psychological data to explanations falling well outside what has been directly observed (4, 5, 6, 17, 18). Certain features of the responses themselves may also yield information on the nature of reduced performance by subjects in an abnormal state, and it is the purpose of this experiment to direct attention to one such possibility. As the latency of the reaction-time response is known to vary with stimulating intensity (2, 3, 12), it is possible to compare the relative variation induced by changes in an identical stimulus variable for both normal and psychotic subjects with an attention to the details and form of the responses obtained. By comparing not only the changes produced in central tendency and group variation, but also in the patterns of intraindividual variability and the relation of initial standing to subsequent change, some of the similarities and differences in performance may be thrown into relief to yield added information on the nature of the deviant psychomotor action known to be characteristic of schizophrenic response.

### B. METHOD

The test procedure is a classical reaction-time situation, employing pure-tone auditory stimulation and a finger-lift response of the right hand. *S* was

---

\* Received in the Editorial Office on June 15, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

instructed to depress a response key upon receiving a preparatory signal, and to release it instantly upon hearing the stimulus tone. Stimulating tones were delivered by an audiometer,<sup>1</sup> through an electronic switch with rapid rise-decay characteristics, to a cushioned headphone on the right ear. The preparatory signal ("ready") was delivered, by modulated voice, to the corresponding left cushioned headphone, preceding the stimulus tone by an interval varying randomly between one and four seconds. The response key was placed in a midline position before *S* on a table 25 inches in height; key tension was maintained at 200 gms. with a full-opening traverse of 1 mm.

The latency of response was recorded by means of a millisecond chronometer.<sup>2</sup> The timer, actuated simultaneously with the onset of a stimulus tone, was arrested by the lift-response of *S* breaking an electric circuit through the response key. All recording apparatus and the experimenter were placed outside the sound-deadened chamber in which *S* sat while making his responses. Each *S* performed 30 practice trials (unscored) to stimuli of medium intensity to acquaint him thoroughly with the test procedure and to achieve stability in response prior to the recording of the experimental measures (5, pp. 64-65, 110).

Three levels of stimulus strength were employed for all *Ss*; intensities of 25, 35 and 50 decibels above the audiometrically established zero-point for a frequency of 1000 cps. Testing always proceeded from stimulation at the lower intensity to the higher. Rest pauses were observed regularly between series of trials at each stimulating intensity, during which the experimenter entered the test chamber for comment and conversation with *S*. Fourteen trials were made at each level of stimulating intensity, the first four of which were universally discarded as a familiarization period and performance on the remaining ten trials was then scored as a measure of reaction latency.

### C. SUBJECTS

Reaction-time latencies to stimuli of varied intensity have been observed for two *S* groups:

*Normal Group.* A sample of 30 non-patients (16 male, 14 female), self-supporting individuals with no history of psychiatric disorder. The age range extended from 21 to 63 years, with an average age of 26.1 years (9).

*Schizophrenic Group.* A sample of 30 patients (15 male, 15 female) with longstanding diagnoses of schizophrenia. The average duration of illness was

<sup>1</sup> Maico, D-10

<sup>2</sup> Standard Electric, MST-500



16.6 years and the group included no patient with less than three years of illness under a diagnosis of schizophrenia. None of the patients was receiving tranquilizing medication at the time of testing, nor had any received such medication for a minimal period of two months before taking part in this experiment. The age range was from 27 to 58 years, with an average of 47.6 years. When judged by such outside criteria as behavior ratings, psychomotor evaluation and Porteus Qualitative scores, the sample studied falls near the middle of the range of disordered behavior usually found among chronic schizophrenic populations (9).

#### D. RESULTS

For convenience of reference, the latencies obtained at the lowest stimulus intensity will be referred to as the "initial standing." Table 1 presents the mean latency and standard deviation for each *S* group for the three levels of stimulating intensity. It is apparent, by inspection, that a different order of latency characterizes response by the two groups, as would be expected on the basis of previous investigation.

TABLE 1  
REACTION-TIME LATENCY (IN MS.) FOR NORMAL AND SCHIZOPHRENIC *SS*  
AT VARIED LEVELS OF AUDITORY STIMULUS INTENSITY

Stimulus intensity (in decibels)	Normal group ( <i>n</i> = 30)		Schizophrenic group ( <i>n</i> = 30)	
	<i>M</i>	$\sigma$	<i>M</i>	$\sigma$
25	340	62.4	931	395
35	247	45.2	623	277
50	213	36.1	480	205

Both groups show successive decreases in average latency with increases in stimulus strength. It may also be noted that group variation, as expressed by the standard deviation, is reduced for both normal and schizophrenic performance with increases in stimulus strength. The relative reduction in mean latency by each group is illustrated, in Figure 1, by a plot in which performance at each level of stimulating strength has been connected by a traced line for ease in reading. The fall in the curve for schizophrenic response appears to be more abrupt, and several correlations have been computed to examine this effect more precisely. Performance by the normal group yields a correlation of .52 ( $P < .01$ ) between latency scores on trials at the 25 and 50 db. levels; the same correlation for the schizophrenic group being .84 ( $P < .01$ ). A second set of relationships, made between performance on initial standing and change with the variable (i.e., latency at 25 db. correlated

with latency at 25 db. minus latency at 50 db.), yields a correlation of .75 ( $P < .01$ ) for the normal group and .68 ( $P < .01$ ) for the patients.

One further statement about relative performance by the  $S$  groups may be made in terms of the intraindividual variability in latency at each level of stimulus strength. As data of this type are less usual, some explanation is required. Each response, by a given  $S$  at a given stimulus intensity, may differ

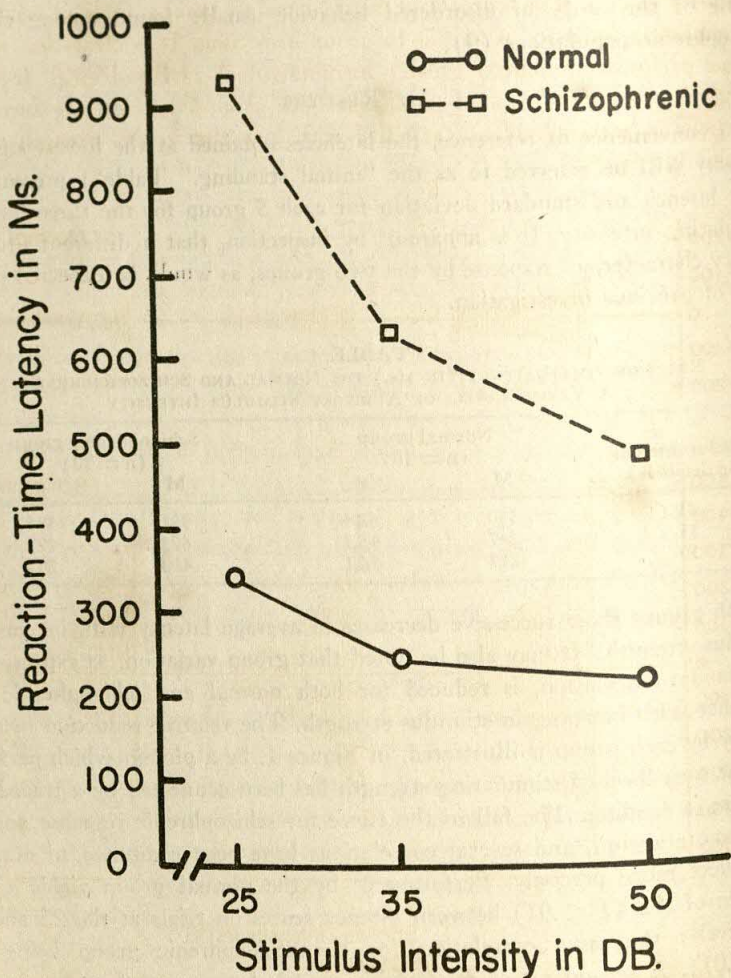


FIGURE 1  
AVERAGE REACTION-TIME LATENCY AT VARIED LEVELS OF STIMULATING INTENSITY



somewhat from other responses made by him to stimuli of the same intensity. If we derive a standard deviation of performance by each *S*, calculated about his own mean for that level and treating his individual performance as a universe, some idea may be had of intraindividual scatter or commonality of reaction. A standard deviation has been computed for each *S* in both groups at all levels of stimulating intensity in this way. The data are shown graphically in Figure 2, in which the length of an individual line, drawn out positively and negatively from zero, represents  $\pm 1\sigma$  of intraindividual variability for each *S*. The scores have been placed in ranked order at each level of stimulating intensity as an aid to group inspection. It is clear that schizophrenic performance reflects greater intraindividual variation at all levels of testing than does the normal, and that both groups demonstrate successive reductions in intraindividual variability with increases in the strength of the stimulus intensity.

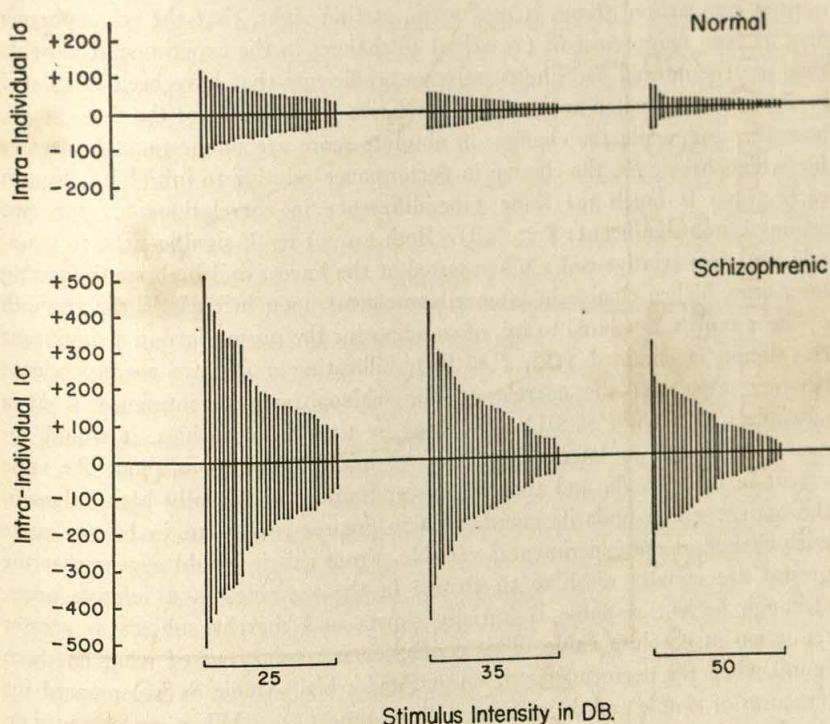


FIGURE 2

INTRA-INDIVIDUAL VARIABILITY AT VARIED LEVELS OF STIMULATING INTENSITY

## E. DISCUSSION

As it is not the purpose of this experiment to stress the differences in latency of reaction-time by the normal and schizophrenic groups as such, no detailed comment or contrasting of group performance will be made, apart from noting that, once again, such differences are readily demonstrable. Attention should be turned, rather, to the fact that both groups are reactive to changes in the experimental variable, and that the responses obtained at differing levels of stimulus strength exhibit systematic and similar trends by each group in terms of average latency and in both group and intraindividual variability.

The reduction of latency to stimuli of increased strength by the normal group is of an order with that previously reported (2), and here forms a standard with which schizophrenic performance may be compared. As the response curve of the patients appears to fall more abruptly than that of the normal comparison group, it may seem, at first sight, that the schizophrenic may be even more sensitive (reactive) to changes in the experimental variable than are the normal *Ss*. The correlation coefficients that have been calculated permit a more detailed analysis of the relative performance of the two groups, however, and while the changes in absolute score are unquestionably greater for schizophrenic *Ss*, the change in performance relative to initial standing in each group is much the same (the difference in correlations for the two groups is non-significant;  $P > .60$ ). Both groups tend, significantly, to maintain the same relative order when tested at the lowest and highest stimulating intensities. When change in latency, consequent upon increase in the strength of the stimulus, is related to initial standing for the normal group a significant correlation is obtained ( $.75, P < .01$ ), indicating a definite correspondence between the two; the correlation for schizophrenic performance is of a similar order ( $.68, P < .01$ ). To judge by these relationships, it would appear that change in latency is relative to initial standing to about the same extent in each group, and that the larger number of initially high values in the patient group finds its counterpart in greater reductions in latency score with change in the experimental variable. From this it would appear that the groups are equally sensitive to change in the variable, in a relative sense, although patient response is initially slower and thereby subject to greater reduction in absolute value. This conclusion is reminiscent of what has been found when the performance of normal and schizophrenic *Ss* is compared for measures of simple and disjunctive reaction-time (1). "When performance on these two types of response is compared, with the simple reaction serving as



a base, the element of discrimination constitutes an equal complication for the chronic schizophrenic group and for the normal" (5, p. 77). It is in contrast to the position reached by Tizard and Venables in their search for a "paradoxical reversal," or re-increase of response latency by schizophrenic *Ss* as the stimulus reaches high strength (14). With the added information that a normal comparison group maintains the same relative correspondence between initial standing and change with the experimental variable, however, and as the data themselves are in agreement between their study and those of this experiment, the point of difference in interpretation is not a sharp one.

The special effort made to permit a closer inspection of patterns of intra-individual variation has been attempted here for the purpose of describing the events of this experiment more completely, and because it is a view of response often needed but infrequently developed when discussing performance of the general type. The display of intraindividual variability coefficients, shown in Figure 2, indicates that the response shapes or "sharpens," regularly, for both normal and schizophrenic *Ss* with an increase in the strength of stimulating intensity. That is to say, each *S* scatters less in response, each response being more like the average, as the stimulus becomes more intense. The reduction in average latency and the shrinkage of intraindividual variability appear to reflect different but related aspects of the vigor of response. As both qualities of response may change at the same time, some degree of correlation will occur between them. This double effect has, at times, called into question the use of parametric statistical techniques for establishing differences between the means of group performances that require these statistics to be uncorrelated. It is often the form taken by the responses themselves, however, and the methods adopted for the analysis of the data must allow for such simultaneous changes in the nature of response. A reduction of latency and a less variable scatter among individual trials may characterize the more vigorous psychomotor response, just as latency and force of contraction may both be qualities of the vigor of a simple reflex action. Another consideration may also be examined with the intraindividual variability data of the present experiment taken as a sample of a situation often met in working with scores derived from psychophysiologic measurement. Where structural limits are placed upon the latency of response, as, for example, in the time required for end-organ resolution or neural conduction, any influence acting to produce increased variation in response will indirectly tend to slow the time of an average response calculated from such data. That is to say, where a variation is more expressible in one direction (slowing) than in another (quickenings), any measure of central tendency computed from a series of more varied

responses will tend in the direction of permissible variation. That the increased latencies of schizophrenic reaction times are not merely so produced may be determined by noting that such increases are also reported by those investigators observing but two or three trials (15), not leading to an averaged measure based upon an elongated one-tailed distribution of individual response. Again, if we select from the data of the present experiment but a single numbered trial at medium stimulus intensity as a measure, e.g., Trial 5, the group mean values calculated at 244 ms. for the normal *Ss* and 595 ms. for the schizophrenic *Ss* are not unlike the values obtained that are based upon a wider sampling of individual performance.

It would appear from these data that while the response of schizophrenic *Ss* in a simple lift reaction-time situation is characteristically increased in latency beyond that of the normal, the psychotic *Ss* are equally responsive, in a relative sense, to changes in the stimulating intensity. Their responsiveness to changes in this physical parameter do not support the idea that the typical slowing of psychomotor reactions in schizophrenia may be attributed to a mere absence of effort or that the state of their behavior disorder is such that all responses to signals from without become diffuse or equally retarded. Rather than implying that the environment is a place from which they receive only cloudy or uncertain signals, the relative similarity in change of response concomitant with stimulus strength, within a defined context, would indicate that their psychomotor behavior, although less sharply focused, remains systematic and predictable by the same factors which govern the behavior of the normal *S*.

#### F. SUMMARY

Simple lift reaction-time responses to stimuli of varied intensity have been observed among normal and schizophrenic *Ss* to examine features of the typically retarded psychomotor response of psychotic *Ss*. Thirty normal and 30 chronic schizophrenics performed RT measures at 3 levels of auditory stimulation following extended practice. The observation of a marked discrepancy in group mean latency and variation was replicated. Both groups showed decreases in latency and intraindividual variability to stimuli of greater intensity, being reactive to the same relative extent. The greater absolute reduction in latency by psychotic *Ss* appears to be a consequent of the initially longer latencies of this group. The findings are related to the ratio of simple to disjunctive RT among comparable subject groups.



## REFERENCES

1. BENTON, A. L., JENTICH, R. C., & WALKER, H. J. Simple and choice reaction times in schizophrenia. *A.M.A. Arch. Neurol. Psychiat.*, 1959, **81**, 375-376.
2. CHOCHOLLE, R. Variation des temps de réaction auditifs en fonction de l'intensité à diverses fréquences. *L'Année Psychol.*, 1940-1941, **41-42**, 65-124.
3. HULL, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67-76.
4. HUSTON, P. E., SHAKOW, D., & RIDGA, L. A. Studies of motor function in schizophrenia: II. Reaction time. *J. Gen. Psychol.*, 1937, **16**, 19-32.
5. KING, H. E. *Psychomotor Aspects of Mental Disease*. Cambridge: Harvard Univ. Press, 1954.
6. ———. Psychomotor tests. Ch. 13 in *Experimental Abnormal Psychology*. Zubin, J., et al., Columbia Univ. Bookstore, 1957.
7. ———. Some explorations in psychomotility. *Psychiat. Res. Reports*, 1961, **14**, 62-86.
8. ———. Psychological effects of excitation in the limbic system. Ch. 31 in *Electrical Stimulation of the Brain*. D. E. Sheer (Ed.). Austin: Univ. Texas Press, 1961.
9. ———. Anticipatory behavior: Temporal matching by normal and psychotic subjects. *J. Psychol.*, 1962, **53**, 425-440.
10. ———. Psychomotor indicators of behavior disorder arising from neurologic trauma and disease. *Psychiat. Commun.*, 1962, **4**, 31-35.
11. OBERSTEINER, H. Über eine neue einfache Methode zur Bestimmung der psychischen Leistungsfähigkeit des Gehirnes Geisteskranker. *Firchow's Arch.*, 1874, **59**, 427-458.
12. PIÉRON, H. De la loi de variation des temps de latence en fonction des intensités excitatrices pour les sensations auditives. *Compt. Rendus Soc. Biol.*, 1919, **82**, 1116-1118.
13. SAUNDERS, E. B., & ISAACS, S. Tests of reaction time and motor inhibition in the psychoses. *Amer. J. Psychiat.*, 1929, **9**, 79-112.
14. VENABLES, P. H., & TIZARD, J. The effect of auditory stimulus intensity on the reaction time of schizophrenics. *J. Ment. Sci.*, 1958, **104**, 1160-1164.
15. WEAVER, L. A. Psychomotor performance of clinically differentiated chronic schizophrenics. *Percept. Mot. Skills*, 1961, **12**, 27-33.
16. WELLS, F. L., & KELLY, C. M. The simple reaction in psychosis. *Amer. J. Psychiat.*, 1922, **2**, 53-59.
17. WULFECK, W. H. Motor function in the mentally disordered: I. A comparative investigation of motor function in psychotics, psychoneurotics and normals. *Psychol. Rec.*, 1941, **4**, 271-323.
18. ———. Motor function in the mentally disordered. II. The relation of muscle tension to the performance of motor tasks. *Psychol. Rec.*, 1941, **4**, 326-348.

Western Psychiatric Institute & Clinic  
 3811 O'Hara Street  
 Pittsburgh 13, Pennsylvania

## PREDICTABILITY OF MULTIDIMENSIONAL ABSOLUTE IDENTIFICATIONS FROM INFORMATION TRANSMITTED WITH UNIDIMENSIONAL STIMULI\*<sup>1</sup>

*Human Factors Research Department, Lockheed-Georgia Company*

GLENN R. HAWKES

### A. INTRODUCTION

Alluisi (1) has pointed out that the amount of information transmitted in absolute discriminations will tend to be maximized when there is (a) equal perceptual spacing of the stimuli; (b) an equal number of *S* and *R* categories; (c) immediate knowledge of results; and (d) a maximum physical range of stimulation. Inasmuch as varying the physical range will serve to limit the perceptual spacing possible with a given number of stimuli, conditions (a) and (d) above may be combined by use of the phrase "optimal spacing of the stimuli." Furthermore, variation of a parameter in the stimulus series other than that being judged can affect information transmitted (5). In addition, experience of *O* has been shown to be of significance either with immediate knowledge of results (4) or without such feedback (5), and individual differences may be quite marked (1).

For many unidimensional absolute identifications, information transmitted ( $I_t$ ) typically falls between 2.0 and 3.0 bits (10). Values less than 2.0 bits have been found for identifications of cutaneous stimuli varying in stridence<sup>2</sup> (subjective intensity of the tactual sensations) or in duration, even under the optimal conditions specified by Alluisi. Hawkes and Warm (7) found in the identification of electrical stridence levels, that  $I_t$  reached a maximum of about 1.7 bits when *O*s were furnished immediate knowledge of results. Although the gain in  $I_t$  produced by providing knowledge of results was not statistically

\* Received in the Editorial Office on June 15, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Data on which this paper is based were recorded at U.S. Army Medical Research Laboratory, Fort Knox, Kentucky.

<sup>2</sup> The word *stridence* was chosen to designate the subjective intensity of tactual sensations for the following reasons: (a) it refers in part to loudness (subjective intensity of acoustic stimulation); (b) it also refers to roughness (a tactual sensation), a word which has been used to describe auditory sensations; (c) its use has been relatively infrequent, but not rare; and (d) its sight and sound is harmonious with analogues in vision and audition; namely, brightness and loudness.



significant in that study, the use of a physical range of stimulation including pain-eliciting intensity levels did yield significantly better performance than restriction of the range to lower levels. In the restricted range situation,  $I_t$  reached a maximum of only 1.4 bits. Absolute identification of electrical cutaneous stimulus duration was found by Hawkes (5) to result in  $I_t$  of up to 1.9 bits. If only stimuli of low stridence were used, however, duration identifications became much more difficult, especially for inexperienced  $O$ s. In agreement with the results for identification of stridence levels, use of a relatively large physical range of stimulation resulted in significantly better performance than the use of a more restricted range. Relatively minor changes in perceptual spacing engendered by separation of the stimuli by an equal number of jnd's or by equal apparent sensation magnitudes made no significant difference in information transmitted.

Many investigators have reported that far more than 3.0 bits may be transmitted if additional dimensions are used in the stimulus display.  $I_t$  with multidimensional identifications has been reported to range from approximately 3.0 bits when identifying combinations of pitch and loudness (11) to a very large number of bits when attending to speech, music, and other complex displays (2).

The present study investigated the channel capacity for multidimensional identifications of electrical cutaneous stimuli varying in both stridence and duration. Based on the reduction in efficiency found when identifying stimulus durations presented at a weak rather than a strong stridence level, particularly if inexperienced  $O$ s are used (5), it was predicted that performance in the multidimensional situation would not be a simple summation of  $I_t$  with these dimensions used singly. Both experienced and naive  $O$ s were used.

## B. METHOD

The apparatus was essentially the same as that used previously (5). Alternating current at a frequency of 100 cps was applied to  $O$  through a 12-mm. diameter electrode resting on the index finger pad; a 25-mm. diameter inactive electrode rested on the palm. The intensity values of the stimuli were the same as the extended range values of Hawkes and Warm (7), i.e., intensity levels eliciting pain were included. Stimulus durations were the same as the extended values used by Hawkes (5), 50 to 1500 msec.

Group 1 consisted of five  $O$ s, all of whom had participated in two different experiments involving absolute identification of electrical cutaneous stimulus duration (5). A second group of five  $O$ s had had no prior experience in studies involving absolute identifications. Both groups participated in all

phases of the present investigation, and all *O*s were given a familiarization session in which electrical cutaneous stimuli were demonstrated in order to avoid any possible complication from the use of intensity levels eliciting pain. No emotional or other difficulties were reported.

In three sessions, identifications were made of stimulus combinations of equal numbers of stridence levels and durations ( $2 \times 2$ ,  $3 \times 3$ , or  $4 \times 4$ ). An additional six sessions were devoted to identification of stimulus combinations with either more stridence levels than durations, or the reverse ( $2 \times 3$ ,  $3 \times 2$ ,  $2 \times 4$ ,  $4 \times 2$ ,  $3 \times 4$ ,  $4 \times 3$ ). Order of presentation was counter-balanced. In order to determine whether or not errors in the  $4 \times 4$  session could be due to relatively close (perceptual) spacing of the stimuli, in an additional session *O*s again identified combinations of two stridencies and two durations. In this tenth session, the stimulus values were the same as the two middle values used in the  $4 \times 4$  session. In each session, *O* was given a chart showing the number of stridence levels and durations to be used in the particular session, and the code number to be used to identify appropriately each combination. All stimulus combinations used in a given session were demonstrated prior to absolute identification by *O*s.

### C. RESULTS

Figure 1 presents mean amounts of information transmitted ( $I_t$ ) for those sessions involving equal numbers of stridencies and durations. Note that none of the *O*s made errors when presented combinations of two stridence levels and two durations. Further, none of the *O*s made errors in the  $2 \times 2$  session in which the stimulus values were the same as the two middle values of the  $4 \times 4$  session. The data on which Figure 1 is based were tested by analysis of variance to determine the influence of Complexity (amount of information available to *O*), Experience, and the interaction of  $C \times E$ . Of these, only Complexity proved to be statistically significant ( $P < .001$ ). The maximum information transmission for these sessions was nearly 3.0 bits.

Mean amounts of information transmitted in sessions that used unequal numbers of stridence levels and durations are presented in Figure 2. There appears to be a tendency for experienced *O*s to transmit more information than naive ones, and for performance to be more efficient for identifications of stimulus combinations with more durations ( $W_D$ ) than levels of stridence ( $W_I$ ). The data of Figures 1 and 2 are combined in Figure 3 to illustrate general trends.

The results of an analysis of variance of the data of Figure 2 are summarized in Table 1; it is apparent that both the experience of *O* and the



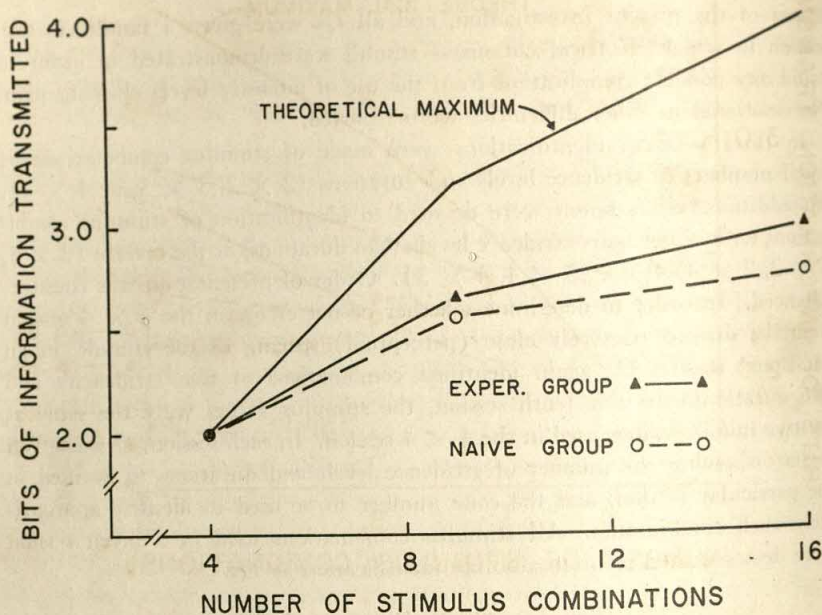


FIGURE 1

MEAN INFORMATION TRANSMITTED WHEN IDENTIFYING COMBINATIONS  
WITH AN EQUAL NUMBER OF STIMULUS INTENSITIES AND DURATIONS

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF INFORMATION TRANSMITTED  
WHEN IDENTIFYING STIMULUS COMBINATIONS CONTAINING  
AN UNEQUAL NUMBER OF DURATIONS AND  
STRIDENCE LEVELS

Source	df	MS	Error term	F	P
1. Subjects (S)	8	855.33	—	—	—
2. Complexity (C)	2	7264.27	13	36.93	< .001
3. Experience (E)	1	7020.02	1	8.21	< .05
4. Weighting (W)	1	5980.02	12	27.06	< .001
5. C × E	2	920.47	11	—	—
6. C × W	2	59.27	10	—	—
7. E × W	1	3.75	11	—	—
8. S × C	16	155.53	10	—	—
9. S × W	8	187.13	10	—	—
10. S × C × W	16	237.93	—	—	—
11. E × C × W	2	788.60	—	—	—
<i>Pooled Error Terms</i>					
12. 10 + 9	24	221.00	—	—	—
13. 10 + 8	32	196.73	—	—	—

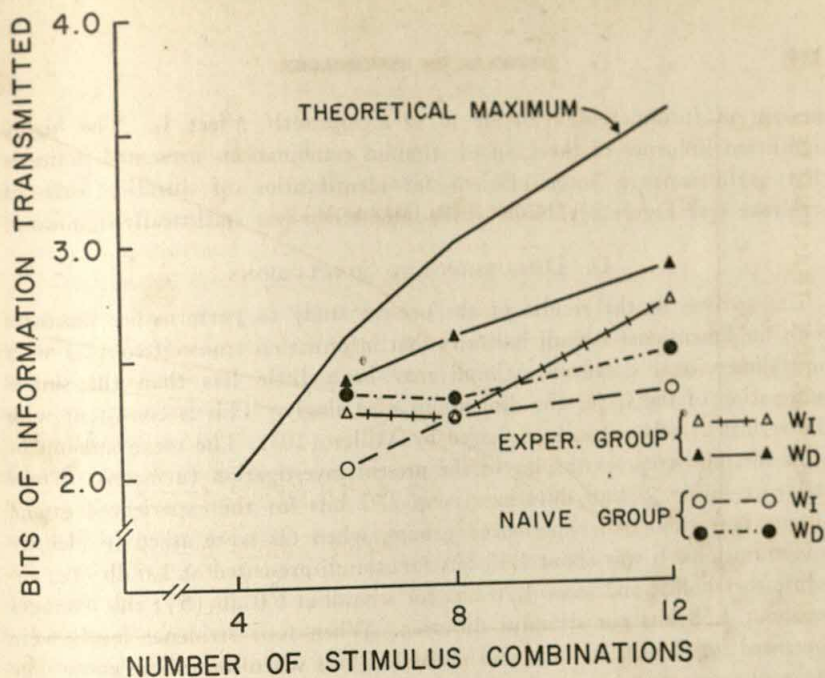


FIGURE 2  
MEAN INFORMATION TRANSMITTED WHEN IDENTIFYING COMBINATIONS  
WITH EITHER MORE STIMULUS INTENSITIES THAN  
DURATIONS ( $W_I$ ) OR THE REVERSE ( $W_D$ )

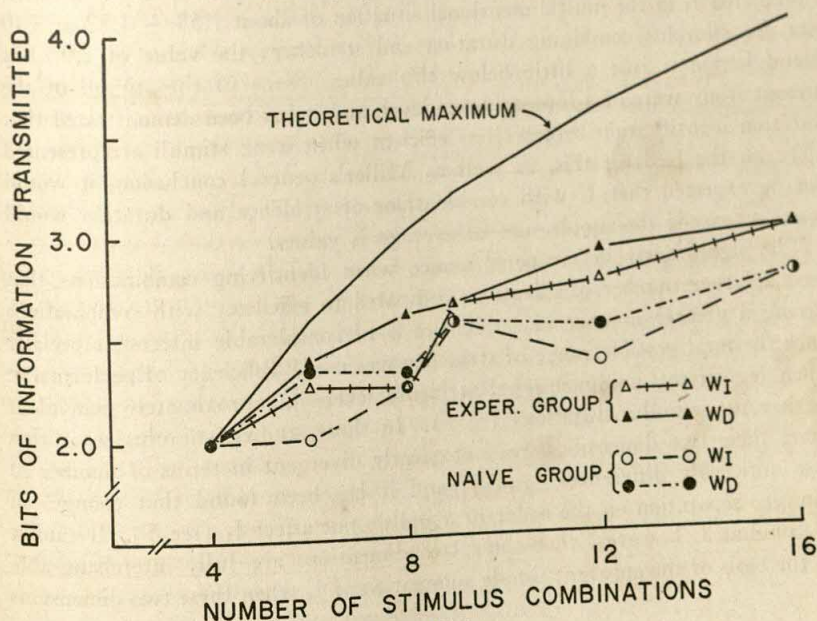


FIGURE 3  
MEAN INFORMATION TRANSMITTED AS A FUNCTION OF THE COMPLEXITY  
AND TYPE OF STIMULUS DISPLAY



amount of information available to *O* significantly affect  $I_t$ . The highly significant influence of the kind of stimulus combinations presented indicates that performance is more efficient for identification of duration than of stridence (see Figure 2). None of the interactions was statistically significant.

#### D. DISCUSSION AND CONCLUSIONS

Comparison of the results of the present study to performance measures with unidimensional stimuli indicates that information transmitted ( $I_t$ ) with multidimensional cutaneous stimuli may be a little less than the simple summation of the  $I_t$  for the dimensions used singly. This is consistent with the results of other studies reviewed by Miller (10). The mean amount of  $I_t$  in the most complex display of the present investigation (a matrix of four stridence levels  $\times$  four durations) was 2.97 bits for the experienced group. In previous studies of experienced groups, when *O*s were asked to identify four durations,  $I_t$  was about 1.45 bits for stimuli presented at 1.6 db (re: absolute threshold), and some 1.70 bits for stimuli at 6.0 db (5); this averages to about 1.58 bits per stimulus duration. When four stridence levels were presented for identification,  $I_t$  was about 1.60 bits when immediate knowledge of results was furnished, and 1.45 bits without such knowledge, and the difference was not statistically significant. This averages to about 1.52 bits per stimulus stridence. The simple addition of these values, then, furnishes a predicted  $I_t$  in the multidimensional situation of about  $1.58 + 1.52$  or 3.10 bits per stimulus combining duration and stridence; the value of 2.97 bits found herein is just a little below this value. Some of the stimuli of the present study were of a low stridence level, and it has been demonstrated that duration identification becomes less efficient when weak stimuli are presented (5); on the basis of this, as well as Miller's general conclusion, it would not be expected that  $I_t$  with combinations of stridence and duration would equal or exceed the simple sum of average  $I_t$  values.

The significantly better performance when identifying combinations that used a larger number of durations compared to efficiency with combinations having a greater number of stridencies is of considerable interest, especially since the total possible range of stridency was used. Efficiency of performance when identifying unidimensional stridence levels is approximately equivalent to that for stimulus durations (5, 7). In those studies, stimulus separation along these two dimensions was not greatly divergent in terms of number of just noticeable differences (jnd's), and it has been found that changes in stimulus separation on the order of a jnd do not affect  $I_t$  (see 5). It cannot be concluded, however, that these two dimensions are fully interchangeable on the basis of the apparent simple summation of  $I_t$  when these two dimensions

are combined, as discussed in the preceding paragraph. The number of jnd's for stridence [about 26 or so according to Hawkes (6)] is perhaps quite limited when compared to the possible jnd's for duration (cf. 8), where a vastly greater physical range of stimulation may be used.

It is suggested that when *O*s are identifying combinations with opportunity for a choice in attention between duration and stridence, detection of duration is easier and thus becomes the preferred dimension. Békésy (3) has reported that there is a variation in apparent area of cutaneous sensation as stimulus intensity is changed. Further, some of the intensity levels of the present study elicited sensations of pain plus tingle, and others only tingle. It is possible that some of these sensational variations may have affected identifications in the present experiment, with the variations being more closely correlated with stridence than duration. McGrath (9) has reported that when attention sharing is required between two modalities in a vigilance task, detection of easy signals improves while detection of difficult signals is unaffected or declines. A similar mechanism might be operative when *O*s are asked to identify cutaneous stimulus combinations with an unequal number of stridence levels and durations.

Békésy also has pointed out many similarities between the operation of the auditory and cutaneous nervous systems. Note in this connection that although  $I_t$  for identifications of unidimensional stridences or durations is less than comparable performance with acoustical stimuli, efficiency with stridence-duration combinations appears to be about the same as reported by Pollack (11) for identifications of patterns of pitch and loudness, about 3.0 bits. In spite of (or perhaps because of) the relatively primitive state of development of the cutaneous nervous system, in some situations it may be about as efficient as the auditory apparatus.

This study indicates that the *kind* of stimulus display used (in this case, use of more durations than stridence levels in the combinations presented) can significantly increase the amount of  $I_t$ . A previous study (5) indicated that the perceived stridence can affect  $I_t$  with duration identifications. In previous studies of cutaneous identifications, neither immediate knowledge of results nor relatively minor changes in stimulus spacing significantly affected transmission of information, whereas relatively gross changes in stimulus spacing resulting from manipulation of the physical range of stimulation did significantly affect channel capacity.

In addition to the conditions specified by Alluisi (1) as influencing channel capacity in a particular study, it may be concluded that experience in tasks of similar nature will have a significant effect both in multidimensional (as in the present study) and unidimensional absolute identifications (4, 5).



## E. SUMMARY

The number of possible absolute identifications of electrical cutaneous stimuli varying in both stridence and duration was investigated for both experienced and naive *O*s. In some sessions, an equal number of stridencies and durations were used; in others there were more stridence levels than durations or the reverse.

Maximum transmission of information was with combinations of four stridence levels and four durations. The channel capacity was about 2.97 bits for the experienced *O*s and 2.76 bits for the naive ones. Experience was not a statistically significant factor when an equal number of durations and stridence levels were used. In sessions with unequal numbers of durations or stridencies, experienced *O*s transmitted significantly more information and both groups were more efficient when identifying combinations with more durations than stridence levels.

Experienced *O*s identifying stimuli varying in stridence or duration under conditions comparable to those of the present study have been reported to transmit about 1.45 to 1.7 bits in each of these (unidimensional) situations. The value of 2.97 bits of this study was just a little less than the simple sum of the performances obtained with the dimensions used singly, but it was also apparent that the dimensions are not fully interchangeable.

## REFERENCES

1. ALLUISE, E. A. Conditions affecting the amount of information in absolute judgments. *Psychol. Rev.*, 1957, **64**, 97-103.
2. ATTNEAVE, F. Applications of Information Theory to Psychology. New York: Holt, 1959.
3. BÉKÉSY, G. VON. Experiments in Hearing. New York: McGraw-Hill, 1960.
4. HARTMAN, E. B. The influence of practice and pitch-distance between tones on the absolute identification of pitch. *Amer. J. Psychol.*, 1954, **67**, 1-14.
5. HAWKES, G. R. Absolute identification of duration. *Percep. & Mot. Skills*, 1961, **13**, 203-209.
6. ———. Cutaneous discrimination of electrical intensity. *Amer. J. Psychol.*, 1961, **74**, 45-53.
7. HAWKES, G. R., & WARM, J. S. Maximum  $I_t$  for absolute identifications of cutaneous electrical intensity level. *J. Psychol.*, 1960, **49**, 279-288.
8. ———.  $\Delta T$  for electrical cutaneous stimulation. *J. Psychol.*, 1961, **51**, 263-271.
9. MCGRATH, J. J. Performance sharing in dual-mode monitoring. *Human Factors Res., Inc., Tech. Rep.*, 740-1, Los Angeles, Calif., 1962.
10. MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.
11. POLLACK, I. The information of elementary auditory displays. II. *J. Acoust. Soc. Amer.*, 1953, 765-769.

Lockheed Aircraft Corporation  
Georgia Division  
Marietta, Georgia

## INCREASED OCCURRENCE OF EEG ALPHA DURING INCREASED ATTENTION\*

*Perception Laboratory, Veterans Administration Hospital, Bedford, Massachusetts*

---

THOMAS MULHOLLAND AND SYLVIA RUNNALS

---

The suppression of the EEG alpha rhythm following a novel stimulus is a classical demonstration. A sudden increase of attention or a change in stimulation has been associated with alpha suppression (1, 3, 4, 6, 8, 12), as have been attention sets and imagery (10, 11, 20, 22) though there is disagreement about the latter (5, 9, 18). Facilitation of the alpha rhythm during attention set had also been reported (24).

In studies of attention, the effect of internalized attention processes is often confounded with the effect of a stimulus change which promotes the attentive act. This problem was met by utilizing the ability to perceive more frequently certain visual illusions (in accordance with instructions) while viewing a stimulus which rotates in a constant cycle. This control of perception, induced by instruction, persisting within *S*, activated and deactivated by *S* at the appropriate time, required continuing attention.

In these experiments the effect of a continuing attention-set (to control perception) on EEG alpha, was compared with the effect of a recurring complex alerting (responding to a recurring signal by giving a report).

### A. EXPERIMENT 1

#### 1. Method

Six normal men viewed a constantly rotating stimulus, reporting their visual perceptions at five-second intervals in response to an audible click. Simultaneously the electroencephalogram (EEG) was recorded.

The stimulus (a  $2 \times 2$  inch white card mounted on a black rod at its mid-point) was enclosed in a black box. *S*, from a distance of nine feet, looked through a  $3 \times 4$  inch aperture in the front of the box and saw the stimulus moving (33.3 r.p.m.) in front of a homogeneous, dark gray background.

Each *S* viewed the stimulus during four conditions (each four minutes long): *Simple viewing (SV)* or *Open-ended set (OE)*. *S* viewed the object

---

\* Received in the Editorial Office on June 18, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



with no specific attention-set; *Attention-set RCW*. Prior to viewing, *S* was instructed to "try to see" only the veridical perception of an object rotating clockwise; *Attention-set TD*. *S* was instructed to "try to see" only a certain illusion (described later); *Viewing a stationary object*. *S* viewed the stationary stimulus and "reported" a number proceeding backwards from 100. During each condition *S* made a verbal report approximately every five seconds. The sequence of conditions was systematically varied.

The EEG was obtained (using a Grass Model III EEG Machine) from scalp electrodes (bipolar) placed over parietal and occipital areas bilaterally. Recordings were made continuously during each condition and during a preceding two-minute "resting" period (eyes closed). The time of the click signal for *S* to make a verbal report was marked on the EEG manually.

## 2. Analysis

The number of reports of the various perceived motions and the number of .1-second units containing alpha were the basic units analyzed. Because of the small number of *Ss* and the large individual differences each *S* was evaluated separately with simple non-parametric statistics. Forty-eight verbal reports were obtained from each *S* during each condition making a total of 1,152 reports. Reports of perceived motion were classified according to the kind of motion reported.

The EEG records were analyzed as follows. Each portion of the record containing alpha (8-13 c.p.s.) on either or both sides was delineated by an experienced EEG technician (S.R.). This delineation was "blind" without knowledge of particular conditions or *Ss* being analyzed. The delineated EEG alpha was measured by T.M. and T.C. to the nearest 0.1 second for each of four 1-second intervals following each verbal report. Thus, 192 seconds of recording were analyzed in each condition for each *S*. The resulting alpha scores were calculated as per cent-time alpha, i.e., the ratio of time containing alpha to total time  $\times 100$ , rounded off to the nearest integer.

## 3. Results

*a. Reports of perceived motion.* While viewing the moving stimulus each *S* reported a variety of perceived motions: the veridical rotation clockwise (rcw) motion, and illusions: (*a*) rotation counterclockwise (ccw); (*b*) oscillation (osc.)—the object seems to make semi-turns left to right, right to left, etc., seeming not to make complete turns; and (*c*) two dimensional motion (td)—the object seems not to turn, but seems to alternately expand and contract in a frontal plane (2, 15, 16, 17, 19).

For each *S* the frequency of occurrence of each kind of perceived motion was obtained for each of 12 successive groups of four consecutive reports. The differences between attention-set conditions were evaluated (for each kind of perceived motion separately) using Friedman's Chi Square of Ranks (21). Rotation clockwise motion was reported most often during the attention-set *RCW* while two dimensional motion was reported most often during the attention-set *TD*. This result was statistically significant for five of the six *Ss* treated separately and for the group as a whole (see Table 1). The mean frequency of report for the rotation clockwise motion and two

TABLE 1  
MEAN PERCENTAGE OF TOTAL REPORTS OF EACH KIND OF PERCEIVED MOTION

Viewing conditions	Rotation CW	Perceived motion		Two-dimension
		Rotation CCW	Oscillation	
RCW	78	7	5	10
S.V.	37	13	11	40
T.D.	32	4	3	61

dimensional motion for each of 12 successive groups of four consecutive reports during each condition was also evaluated. The results shown in Table 1 were maintained throughout the viewing period.

*b. EEG Alpha.* All *Ss* showed the well-known suppression or reduction of alpha amplitude during the viewing conditions relative to the eyes-closed conditions. Also, alpha occurred more often while viewing a stationary stimulus compared to viewing a moving stimulus. For each condition of viewing a moving stimulus the total number of .1-second units containing alpha was obtained for each *S* separately. If alpha were equally probable during each condition equal parts of the total alpha would be obtained during each condition. Expected alpha was compared with that actually obtained using Chi Square. Four *Ss* produced alpha significantly more often during both attention-sets compared to simple viewing; one *S* produced alpha significantly more often during one attention-set, less during the other, relative to simple viewing. One *S* produced alpha significantly less often during both attention-set conditions. All Chi Square tests were significant beyond the .01 level.

This unexpected finding was examined in detail in terms of the distribution of alpha between and during verbal reports and throughout the viewing period. The average alpha for each one-second interval following each report was obtained for each condition. For purposes of comparison the eyes closed "resting" periods were divided into successive five-second units and the amount of alpha in each one-second unit obtained. These were then pooled over all



five-second units. The relative facilitation of alpha during attention-sets was clearly evident. The increase of alpha following verbal reports was also evident. Friedman's Chi Square of ranks for the difference between successive one-second intervals was 9.4;  $P < 0.05$ . See Table 2.

TABLE 2  
MEAN PER CENT TIME ALPHA FOR EACH ONE SECOND INTERVAL  
FOLLOWING EACH VERBAL REPORT\*

	1st	2nd	Intervals 3rd	4th	$\bar{X}^{**}$
Eyes closed	68	69	72	64	68
Viewing stationary stimulus	32	39	37	34	35
Moving stimulus					
Attention-set TD	21	24	30	30	26
Moving stimulus					
Attention-set RCW	21	25	28	29	26
Moving stimulus					
Simple viewing	13	17	18	22	17
$\bar{X}$ (viewing conditions)**	22	26	28	28	

\* Each percentage based on 288 seconds

\*\* Based on 1,152 seconds

The distribution of alpha around the verbal reports was examined for the first 24 reports. An interval of two seconds (one second before and one second after each report) was divided into 0.2-second units. This interval was pooled over all reports for each condition. There was a facilitation of alpha during the attention-set conditions relative to the simple viewing condition. At the same time the familiar suppression of alpha following the signal to make a verbal report was seen for all conditions.

The familiar, expected increase of alpha as a function of duration of viewing was evaluated by dividing the record into successive units of 16 seconds (four seconds following each of four consecutive reports). The increase of alpha over time was clearly evident. Friedman's Chi Square of Ranks for differences among twelve successive alpha scores (pooled viewing conditions) was 26.5;  $P < 0.01$ .

The increased alpha during conditions of attention-set relative to the simple viewing condition was also evident here. The differences among attention-set and simple viewing conditions were tested by ranking each condition at each of the twelve time intervals for the group data and for each *S* separately. Friedman Chi Square of Ranks for the group data was 12.5;  $P < 0.01$ . Each of four *Ss* showed increased (three significantly) alpha occurrence during both attention-sets relative to the simple viewing condition; one *S*

produced alpha significantly more often during one attention-set, significantly less often during the other; one *S* showed a non-significant decrease of alpha occurrence during both attention sets relative to the simple viewing condition.

In sum, for this group of *Ss* and a large sample of EEG recording, per cent time alpha increased during two conditions of attention-set relative to a condition of simple viewing without a specific attention-set.

## B. EXPERIMENT 2

The previous study, though demonstrating alpha facilitation during certain attention-sets did not permit a generalization to other *Ss* or other attention-set conditions. This second study utilized a larger sample of subjects and attention-set conditions.

Stimulus conditions were so arranged that the control of perceptual illusions in accordance with instructions was more or less difficult. To this end, a stimulus object was selected which generated two unique illusions (17). The occurrence of these illusions depended on the brightness contrast between the stimulus and its background. By instructing *S* to see more of one illusion under illumination conditions which rendered that illusion unlikely, a difficult task was produced; instructions to see more of an illusion under illumination conditions which rendered the illusion very likely produced an easier task. In this case, as illumination increased one illusion occurred more frequently while the other less frequently; as illumination decreased the reverse was true. In this way difficulty could be evaluated independently of illumination level.

### 1. Method

The apparatus and procedure were the same as Experiment 1 except for the stimulus object and the greater variety of viewing conditions. The stimulus object was shaped like a "T" with a disk (white on both sides) at one end of the crossbar and another disk (black on both sides) on the other end. The disks were about .5 inches in diameter and the crossbar about two inches long, .25 inches wide. The reflectance of the white disk was 36 per cent, the black, 3 per cent. The stimulus was viewed in front of three backgrounds: light grey (Munsell 9.0 reflectance 79 per cent); middle grey (Munsell 4.5 reflectance 15 per cent); dark grey (Munsell 3.5 reflectance 9 per cent).

Each of 10 normal men gave 36 reports of their perceptions (at five-second intervals) under each of nine conditions presented in random sequence. Simultaneously the EEG was recorded. The nine conditions were combinations of the three backgrounds with three attention-set conditions: (a) simple viewing; (b) attention-set white-front: *S* was instructed to "try to see" only an



illusion whereby the white disk seems to oscillate in front of the axis of rotation; and (c) attention-set black-front: *S* was instructed to try to see an illusion whereby the black disk seemed to oscillate in front of the axis of rotation. (Both of these illusions will be described later.)

The schedule of recording the EEG was the same for each condition. *S* was first instructed concerning the attention-set to be adopted followed by two-minute period (eyes closed) to obtain a resting EEG. Two seconds before the end of this period a "ready" signal was given and *S* opened his eyes. He then viewed the object giving verbal reports concerning his perceptions every five seconds in response to an audible click.

The EEG was recorded as before. The audible signal to make a verbal report was automatically marked on the EEG record. Each *S* gave 36 reports under each condition yielding a total of 3,240 reports. Four seconds after each report was analyzed for EEG alpha for a total of 12,960 seconds.

## 2. Results

*a. Reports of perceived motion.* A variety of perceived motions was reported. Two illusory motions were unique to this stimulus object (white-front and black-front) while the others (rotation clockwise, rotation counterclockwise and two dimensional motion) were similar to those described in Experiment 1. With these illusions the object seemed to oscillate back and forth making only semi turns. With *white-front* illusion the oscillation motion so occurred that the *white* disk seemed to swing back and forth *in front* of the axis of rotation with the black disk swinging behind. With *black-front* illusion the apparent position of the disks were reversed. Only the white-front and black-front illusions were affected by attention sets and background contrast. The other perceived motions were not affected.

*White-front* occurred more frequently the darker the background while *black-front* occurred less frequently (17). Also, as in Experiment 1, the occurrence of the illusions depended upon attention-set. Table 3 presents the mean percentage of total responses for each illusion for each condition. Table 4 presents the number of *Ss* reporting *white-front* more often than *black-front* during each condition. Because of this replication of previous results and the large numbers of responses sampled the only statistic presented is the percentage of total responses. This effect of attention-sets was also evident throughout the viewing period.

*b. EEG alpha.* Each four-second interval following each report was analyzed for alpha as described in Experiment 1. These were pooled over the ten *Ss* yielding a matrix of 324 mean per cent-time alpha scores, 36 under each of

nine conditions. These means were analyzed using analysis of variance. The variance was divided among attention-sets, backgrounds, first vs. second half of the viewing period and the interactions of these. A summary of the analysis is presented in Table 5.

For this particular sample of EEG records alpha occurrence varied sig-

TABLE 3  
MEAN PERCENTAGE OF TOTAL REPORTS\* FOR EACH ILLUSION

	White-front		Attention-sets Simple viewing Perceived motion		Black-front		$\bar{X}^{**}$	
	wf	bf	wf	bf	wf	bf	wf	bf
Backgrounds								
Dark grey	59	05	40	14	30	29	43	16
Middle grey	51	04	17	12	03	51	24	22
Light grey	34	09	08	21	00	53	14	28
$\bar{X}^{**}$	48	06	22	16	11	44		

\*  $N_1 = 360$  reports

\*\*  $N_2 = 1,080$  reports

TABLE 4  
NUMBER OF SUBJECTS REPORTING WHITE-FRONT MORE OFTEN  
THAN BLACK-FRONT DURING EACH CONDITION

Background	Attention-set	wf > bf	bf > wf
Dark grey	WF	9	1
	SV	9	1
	BF	4	6
Middle grey	WF	8	2
	SV	4	6
	BF	1	9
Light grey	WF	6	4
	SV	2	8
	BF	0	10

TABLE 5  
SUMMARY ANALYSIS OF VARIANCE FOR MEAN ALPHA SCORES

Source	Mean square	df	F	P
Attention-set (A)	451.9	2	6.0	< .01
Backgrounds (B)	405.3	2	5.3	< .01
Half (H)	6552.0	1	86.5	< .001
A × B	747.5	4	9.8	< .001
A × H	330.9	2	4.3	< .05
B × H	3.3	2	< 1.0	> .10
A × B × H	14.7	4	< 1.0	> .10
Residual	75.7	306	—	—
Total	—	323	—	—



nificantly as a complex function of many variables. In general, alpha occurred more often during the second half of the viewing period and occurred somewhat less often during the attention-sets compared to the simple viewing condition. This latter result needs further evaluation in terms of the significant interaction between backgrounds and attention-sets. For the simple viewing conditions alpha occurred *less* often as background reflectance increased. However, for both attention-set conditions alpha occurred *more* often as background reflectance increased.

The distribution of alpha between verbal reports was analyzed in detail. An interval of  $-3$  to  $+2$  seconds around each report was divided into .2-second units. The mean per cent time alpha (expressed here as an estimated probability) was obtained for the group of 10 Ss. In this way an alpha gradient having 25 points each based on 36 seconds of recording was obtained. These alpha gradients are presented in Figure 1. The effect of the complex alerting is clearly evident for all conditions: alpha decreased just before, during and just after a report, increasing between reports. However, the effects of attention-set are not the same for each contrast condition. Alpha increased during attention-set relative to simple viewing with a light background; alpha decreased during one attention-set, increased during the other, with the middle background; alpha decreased during both attention sets with the dark background. There was no evident relation between the occurrence of alpha and task difficulty (as defined previously).

### C. EXPERIMENT 3

This experiment concerned changes of EEG alpha occurrence during a continuing attention-set and during shifts of attention-set.

#### 1. Method

The apparatus and recording of the EEG was the same as in Experiment 1. The procedure for collecting perceptual data and the schedule of attention-sets were different. S viewed the rotating object for an extended period reporting the occurrence of the veridical perception (rotation clockwise) by pressing a signal key as long as the motion was perceived. When S so decided, he shifted his attention-set from simple viewing to a set "to see" only the rotation clockwise motion. Later, when he so decided, he shifted his set back to simple viewing. Later again, he shifted to the rotation clockwise set, then back to simple viewing, etc. In this way S had control over the schedule of set activation.

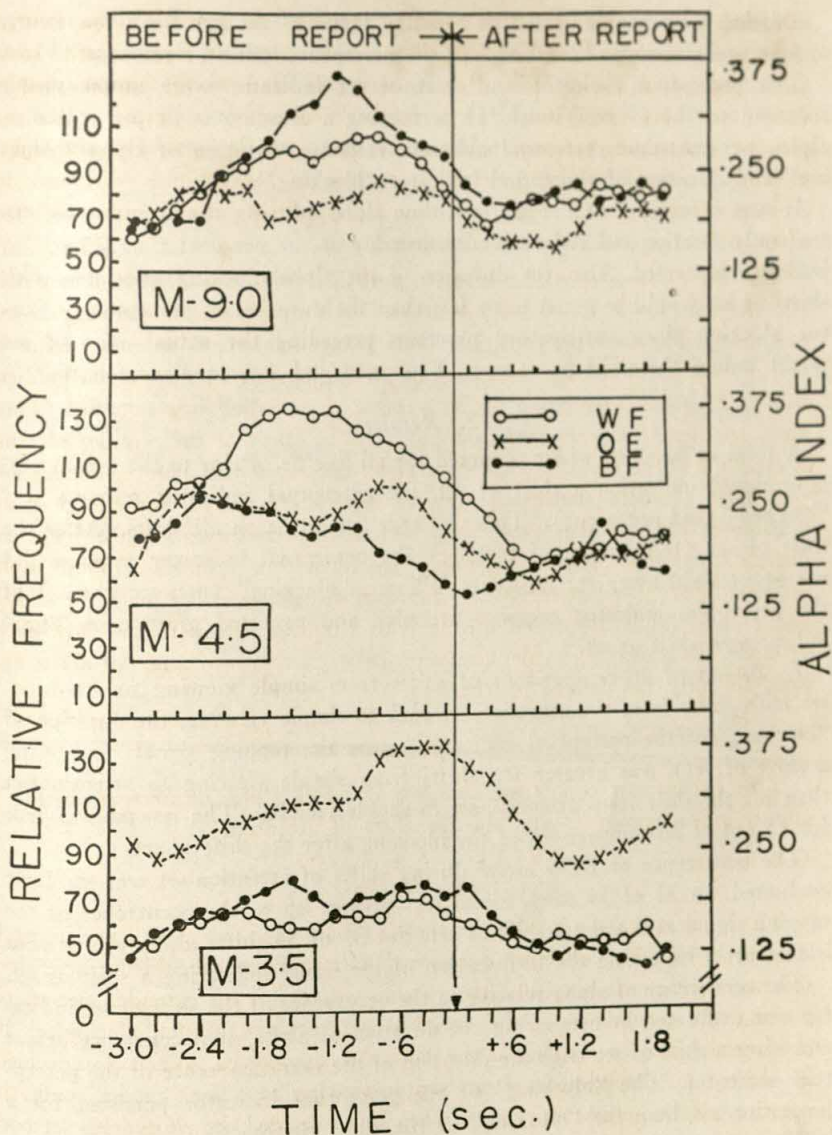


FIGURE 1

Mean gradients of alpha occurrence before, during and after an audible signal to make a verbal report. Alpha occurrence is expressed as a relative frequency of .1 sec. units and as a ratio of relative frequency to the total possible. OE is simple viewing. BF and WF are viewing with a set to see only *ae* and *bf* motions respectively.



Shifting a set was indicated by rapidly "tapping" the key for a few tenths of a second, forcing a "blocking" of the perceptual indicator response.

The perceptual indicator and shift of set indicator were automatically recorded on the Grass Model III permitting a comparison of per cent-time alpha, per cent-time perceptual indicator response, duration of alpha "blocking" and duration of perceptual indicator "blocking."

It was expected that the per cent-time alpha (during the attention-set "to see" only rotation clockwise) would decrease as the per cent time perception indicator increased. Also, the duration of the alpha blocking associated with shifts of set would be equal to or less than the duration of perception indicator blocking since anticipatory processes preceding the actual shift of set might reduce the alpha occurrence. Four men and a woman were tested.

## 2. Results

A total of 52 shifts of set occurred for all five Ss. Prior to the actual tapping signal indicating a shift of set, the perceptual indicator response was "blocked" (did not occur). This was true for all Ss on all shifts. After the shift of set (termination of tapping) the perceptual indicator response did not occur right away for all shifts. Thus, a "blocking" (non-occurrence) of the perceptual indicator response preceded and persisted after every signal indicating a shift of set.

In 20 out of 26 comparisons of shifts from simple viewing to attention-set with shifts from the attention-set back to simple viewing, the duration of "blocking" of the perceptual indicator before the tapping signal (indicating a shift of set) was greater for shifts from simple viewing to attention-set than for the shift from attention-set to simple viewing. The reverse was true for 15 out of 26 comparisons of the blocking after the shift of set.

The occurrence of EEG alpha during shifts of attention-set was similarly evaluated. In 31 of 54 cases, alpha was present up to the occurrence of the tapping signal indicating a shift of set. For 35 of 54 shifts alpha was present immediately following the termination of the signal indicating a shift of set.

The occurrence of alpha relative to the occurrence of the shift of set indicator was evaluated by comparing the duration of alpha non-occurrence before and after a shift of set with the duration of the non-occurrence of the perceptual indicator. The "blocking" of the perceptual indicator persisted for a longer period than the "blocking" of the alpha rhythm.

The mean per cent-time alpha and mean per cent-time perception indicator (both divided by 100) for a period of 10 seconds preceding and 10 seconds following each shift of set were calculated for each .2-second unit. These scores

are presented in Figure 2 for the two directions of shifts: simple viewing to attention-set and the reverse. The systematic variation of the perceptual indicator beginning before the actual shift of set and persisting afterward is clearly evident. However, no such change of EEG alpha was noted, even during the occurrence of the "tapping" signal indicating a shift of set.

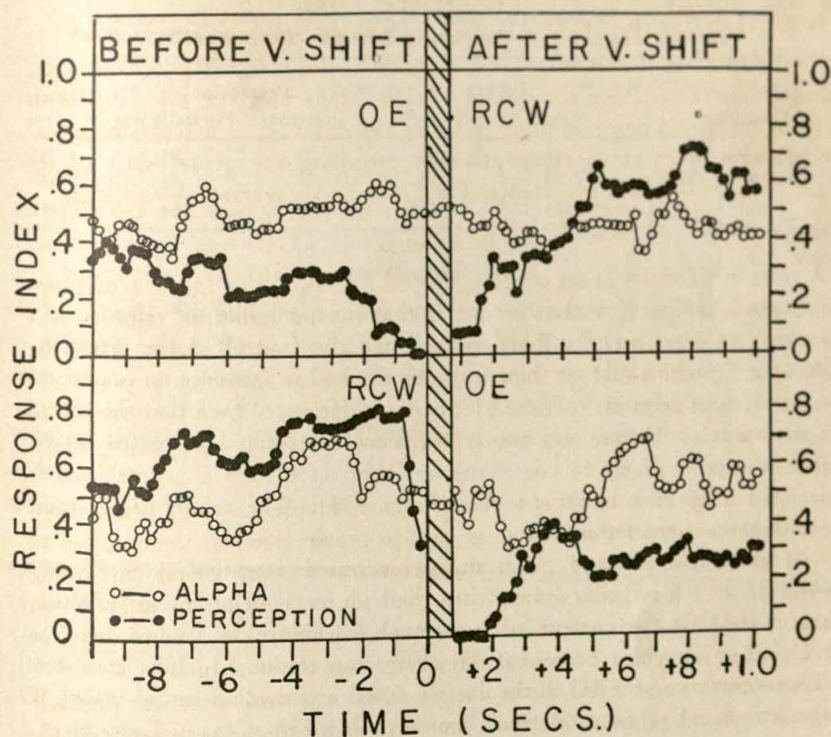


FIGURE 2

Estimated probability of occurrence of EEG alpha and the perceptual indicator response before, during and after a shift of set initiated by the subject. Vertical cross-hatched area in the middle is the average duration of the "tapping" response indicating a shift of set. *OE* is simple viewing; *RCW* is viewing with a set to see only *rcw* motion.

Also apparent (and confirming the previous experiments) is the effect of the attention-set on the key response indicating that *S* was actually perceiving the rotation motion. For all *Ss* when the attention set to see only rotation clockwise was present, the perceptual indicator occurred more often. When simple viewing was present, the perceptual indicator occurred less often.



These results were summarized by the estimated probabilities of occurrence of the alpha and perception indicators for each attention-set condition. See Table 6. Ss divided their time almost equally between simple viewing and the active attention-set. A decrease of alpha during the attention-set was not evident.

TABLE 6  
MEAN ESTIMATED PROBABILITIES FOR EEG ALPHA-NO ALPHA OCCURRENCE AND  
FOR THE INDICATOR OF PERCEIVING ROTATION CLOCKWISE MOTION

Condition <sup>a</sup>	EEG		Perception		Per cent Total time
	Alpha	No alpha	Indicator	No indicator	
Simple viewing	.32	.19	.19	.33	.51
Attention-set RCW	.33	.15	.27	.21	.48
Total	.65	.34	.46	.54	.99

#### D. DISCUSSION

Some years ago K. S. Lashley stated that functions like the *selection of the content of experience* (authors' italics) and the control of the direction of thought; attention and set show a complexity and dependence on organization which is most certainly cortical (13, p. 426). He noted then that the reticular system was so "diffuse and poorly organized for limited patterned activity" that it was "unlikely to contribute anything more than a general undifferentiated facilitation in these processes" (p. 434). The results of the present experiments support this view.

If the suppression of alpha during increased attention is mediated by processes in the reticular system then the high levels of attention reflected by the control of the content of perceptual experience as studied here have little effect on those processes. However, suppression of alpha consistently occurred as Ss responded to the audible signal and made a verbal report. On the other hand, when S "alerted" himself to shift his set (as in Experiment 3) there was no definite, no consistent change of alpha occurrence. In sum, these results reiterate the fact that the term "attention" refers to neither a behavioral nor a neurophysiological process which is a qualitatively consistent entity hence the classical alpha-attention hypothesis so frequently utilized in reference book and review (7, 23) may only refer to a special case of transitory alerting to an external signal.

Alpha suppression was most consistently associated with those alerting processes which were activated by a recurrent *external* signal and which were *transitory*. Those processes showing little relation to alpha suppression were continually activated *from within* (after an initial starting signal) and were

*enduring*. The internal nature of the activation of the attention-sets and of their continuance must require complex cognitive processes involving symbolization, remembrance and volitional initiation. In addition, the perceptual activity being controlled was highly differentiated and organized.

The classical (and still familiar) alpha-attention hypothesis is not qualified in terms of a particular kind of attention process or behavior. The results of these studies show that such a qualification is useful. Such a differentiation of classes of alpha-"attention" relationships may broaden the interpretation of EEG studies relating the alpha rhythms to attention processes.

#### E. SUMMARY

In three experiments the occurrence of EEG alpha during continuing attention-sets and during recurring alerting responses was compared.

Alpha was frequently facilitated during periods of attention-sets or was little affected. On the other hand, the familiar suppressions of alpha occasioned by alerting to an external signal was clearly evident. The behavioral effects of both the attention-sets (on perceptual content) and alerting (giving a verbal report in response to an audible signal) were definite and consistent.

It was concluded that the term attention can refer to neither a qualitatively consistent behavioral nor neurophysiological entity and that the classical and familiar alpha-attention hypothesis refers to a special case, i.e., transitory alerting to an external signal.

#### ACKNOWLEDGEMENTS

The assistance of Thomas Casey (who tabulated much of the data and performed the statistical calculations) and Budd Robb who prepared the Figures is appreciatively acknowledged.

#### REFERENCES

1. ADRIAN, E. D., & MATTHEWS, B. H. C. The Berger rhythm: Potential changes from the occipital lobes in man. *Brain*, 1934, **57**, 354-385.
2. AMES, A. Visual perception and the rotating trapezoidal window. *Psychol. Monog.*, 1951, **65**, No. 7, 3-31.
3. BAGCHI, B. K. The adaptation and variability of response of the human brain rhythm. *J. Psychol.*, 1937, **3**, 463-485.
4. BAKES, F. Effect of response to auditory stimulation on the latent time of blocking of the Berger rhythm. *J. Exp. Psychol.*, 1939, **24**, 406-418.
5. BARRAT, P. E. Use of the EEG in the study of imagery. *Brit. J. Psychol.*, 1956, **47**, 101-114.
6. BERGER, H. Über das Elektrenkephalogram des Menschen. *J. Psychol. u. Neurol.*, 1930, **40**, 160-179.
7. BRAZIER, M. A. B. The development of concepts relating to the electrical activity of the brain. *J. Nervous & Mental Diseases*, 1958, **126**, 303-321.



8. CRUICKSHANK, R. M. Human occipital brain potentials as affected by intensity-duration variables of visual stimulation. *J. Exp. Psychol.*, 1937, **21**, 625-641.
9. DREVER, J. Further observations on the relation between EEG and visual imagery. *Amer. J. Psychol.*, 1958, **71**, 270-276.
10. GOLLA, F., HUTTON, E. L., & WALTER, W. G. The objective study of mental imagery. I. Physiological concomitants. *J. Mental Science*, 1943, **89**, 216-223.
11. HARRISON, J. M. An examination of the varying effect of certain stimuli on the alpha rhythm of a single normal individual. *Brit. J. Psychol.*, 1946, **37**, 20-29.
12. JASPER, H. H., & CRUICKSHANK, R. M. Variations in blocking time of occipital alpha potentials in man as affected by the intensity and duration of light. *Psychol. Bull.*, 1936, **33**, 770-771.
13. LASHLEY, K. S. Dynamic processes in perception. In Adrian, E. D., Bremer, F., & Jasper, H. H. (Eds.), *Brain Mechanisms & Consciousness*. Oxford: Blackwell, 1954, 422-443.
14. MARTINSON, B. M. A study of brain potentials during mental blocking. *J. Exp. Psychol.*, 1939, **24**, 143-156.
15. MILES, W. Movement interpretation of the silhouette of a revolving fan. *Amer. J. Psychol.*, 1931, **43**, 392-405.
16. MULHOLLAND, T. Motion perceived while viewing rotating stimulus-objects. *Amer. J. Psychol.*, 1956, **69**, 96-99.
17. ———. The "swinging disk" illusion. *Amer. J. Psychol.*, 1958, **71**, 375-382.
18. OSWALD, I. The EEG, visual imagery, and attention. *Quart. J. Exp. Psychol.*, 1957, **IX**, 113-118.
19. PASTORE, N. Some remarks on the Ames oscillatory effect. *Psychol. Rev.*, 1952, **59**, 319-323.
20. SHORT, P. L. The objective study of mental imagery. *Brit. J. Psychol.*, 1953, **44**, 38-51.
21. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956, 166-173.
22. WALTER, R. D., & YEAGER, C. L. Visual imagery and electroencephalographic changes. *EEG Clin. Neurophysiol.*, 1956, **8**, 193-199.
23. WALTER, W. G. Intrinsic rhythms of the brain. In Field, J. (Ed.), *Handbook of Physiology*, Section 1: Neurophysiology, (Ed.) Magoun, H. W., Washington: American Physiological Society, 1959, 279-298.
24. WILLIAMS, A. C. Facilitation of the alpha rhythm of the electroencephalogram. *J. Exp. Psychol.*, 1940, **26**, 413-422.

VA Hospital  
200 Springs Road  
Bedford, Massachusetts

PERSONALITY AND ATTITUDINAL CORRELATES  
OF RESPONSE TO DRUG TREATMENT IN PSY-  
CHIATRIC OUTPATIENTS: 1. THEORETICAL  
ORIENTATION AND MEASURING  
INSTRUMENTS TO BE  
EMPLOYED\*<sup>1</sup>

*University of Pennsylvania School of Medicine*

---

K. RICKELS AND R. W. DOWNING

---

A. PURPOSE

The following paper is the first of a series of reports describing the relationships of several personality and attitudinal variables to the responses of neurotic outpatients to treatment with tranquilizing drugs. The factors of treatment response upon which we shall focus our attention are improvement as seen by both patient and doctor, premature termination (dropping out) and the occurrence of side reactions and placebo effects. We shall attempt to present the conceptual framework within which we are operating, discuss the personality and attitudinal effects which we intend to study, and provide background information about the battery of tests we have elected to employ.

B. RELATED STUDIES

A long history of observations on the "placebo effect" (cf. 45) reveals that, in a significant proportion of patients, a wide range of symptoms, ranging from the common cold to varied forms of psychic distress, respond favorably to treatment with an essentially inert and inactive agent. Furthermore, the administration of placebo may result in disturbing side effects (42, 53). Employing the research on the placebo effect as a point of departure, we shall first discuss the importance of personality and attitudinal factors to the understanding of response to drug treatment.

In the first place, it has been demonstrated that response to the pharmacologically inactive placebo has been found to be intimately related to the patient's expectations and to several aspects of his personality organization. Rickels, *et al.* (43) have shown the importance of patient "set" by adminis-

---

\* Received in the Editorial Office on June 18, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Work supported by USPHS Grant MY 2934.



tering to one patient group a placebo given in the same colored capsule for a six-week period, and to a second group placebo in three different forms (two different capsules and one tablet). The first group was led to believe that they were receiving a single drug which would probably help them, but which might have to be taken continuously for a six-week period before improvement would result. The second group was told that they were receiving three distinct drugs (each for a two-week period), one or all of which might prove helpful. A significantly larger number of drop-outs occurred in the former group when there was no hope of medication change for six weeks. Also, despite the fact that physician ratings showed no difference in the amount of improvement for the two groups, the group that believed itself to be receiving several drugs in succession produced higher self-ratings of over-all improvement than did the "same drug" group. Similarly, Kast and Lesch (28) have demonstrated that induced sets concerning the nature of treatment and the efficacy of drugs influence patient responses to psychopharmacological agents. Hence, it seems to us very likely that the patient's expectations concerning the effectiveness of medication may be a large factor in determining his treatment response.

A second important implication of the placebo effect lies in the observation that those patients in whom it may be observed seem to have characteristic personality attributes. Furthermore, there seems to be evidence that the personality characteristics of placebo responders also appear in patients responding with inappropriate side reactions to active agents. Thus, Rickels, Baum and Fales (42) found that side reactions to placebo and to imipramine were both more frequent in neurotic patients attending a general medical clinic than in neurotic patients attending a psychiatric clinic. The former patients were more markedly passive-dependent, more immature, more covertly hostile, more somatically centered, and more frequently oriented toward receiving medicine from an all-powerful and all-knowing doctor than were patients from the psychiatric clinic. Again, von Felsinger, Lasagna and Beecher (16), found that placebo effects and atypical reactions to amphetamine, heroine, morphine, and pentobarbital were more common among their college student volunteer subjects who tended to be immature, impulsive, anxious, hostile, and possessed of a strong, but diffusely directed, drive for achievement. These subjects were emotionally more expressive and had more somatic symptoms during stress. Lasagna, *et al.* (33) found similar patterns in post-operative non-psychotic patients. Joyce (27) observed in a normal student population that placebo reactors, as compared to non-reactors were less critical, more sensitive to social influence, more highly valued by others, and less dominant

and confident; they also showed a greater tendency to report symptoms which might be due to a greater awareness of actual bodily changes.

The common elements of personality structure in these diverse placebo reactor groups (leaving aside those attributes which are probably exclusively characteristic of volunteer groups as such), would seem to be dependency, compliance, immaturity, a high overt or covert hostility level, and predisposition to somatic reactivity.

The third implication of the importance of personality variables concerns drug research methodology. As Beecher (2) has pointed out, response to placebo is typically used as the baseline against which the effects of active agents are evaluated. While the random assignment of subjects to experimental conditions may equally distribute those individuals with personality organizations more prone to side reactions and/or placebo response over the several experimental conditions employed, the differences in responses between such individuals and those with other personality characteristics will inflate experimental error and may very likely detract from apparent drug effectiveness. Also, such a randomization procedure would not permit a study of the interaction of pertinent personality characteristics and drug effects. It may well be that a drug may have effects which differ widely depending upon the personality organization of the patient to whom it is administered (see below) (cf. 35).

Focusing now our attention more directly upon the nature of the interaction between personality factors and response to drugs, there seems to be evidence for an interactional effect both at a neurophysiological and at what might be called a psychological level. Shagass (46) has found that tolerance to sedatives measured in terms of a sedation threshold (or time to reach the point of inflection on the curve relating high frequency EEG wave amplitude and the amount of intravenously injected pentobarbital per kg. of body weight) may vary along continua of: 1. anxiety, 2. hysteria-obsessiveness, and 3. introversion-extroversion. Individuals with greater amounts of anxiety, more pronounced obsessive features, or larger quantities of introversion require larger amounts of the drug before the sedation threshold is reached. Shagass cites evidence that the same effect can be obtained with alcohol and bromides as well as with barbiturates, and hence feels that it is not dependent upon a specific drug. Furthermore, Shagass and Lipowski (47) present evidence that exitant tolerance may co-vary with sedation tolerance, so that the same individuals may be high in both. As Shagass suggests, a possible mechanism serving as the substrate to both introversion and high drug tolerance might be the higher cortical inhibitory potential which Eysenck (15) attributes to introverted and dysthymic individuals.



At the psychological level, it is necessary to consider the intra- and interpersonal significance to the individual patient of the primary drug reaction. After noting that mescaline can result in either satisfaction or apprehension, depending upon personality and situational factors, von Felsinger, *et al.* (16), state, "The affective change may therefore be visualized as not directly determined by the drug but related to the alteration by the drug of such psychic factors as needs, drives, perception, and the change then experienced with pleasure or alarm depending on the total psychic structure." They suggest that, in those individuals, for whom control of hostile impulses is tenuous, the stimulating effect of amphetamine may, as a result of more intensified strivings for control, have a sedating effect.

In a treatment situation where patients with mild anxiety and depression are being treated with tranquilizing drugs, it may well be that neurophysiological and psychological aspects of personality organization tend to reinforce each other in influencing the nature of the reaction to the drug being administered. The work of Shagass would seem to imply that individuals who tend to be more obsessive and more introverted will be less likely to experience a sedating effect from medication than will individuals who tend to be more hysterical and more extroverted. At the same time, however, obsessive patients are often afraid of sedation because they fear the losing of control and therefore they fight the sedative effect and do not become sedated. Also, the rigidity which is a frequent correlate of obsessiveness is likely to render the patient less predisposed to change, and hence make less likely his acceptance and assimilation of symptomatic improvement. Similarly, the suggestibility and possibly the compliance which is frequently associated with hysterical features may lead the patient to experience symptomatic improvement to the extent that he feels that this is what the doctor expects of him. Hence, on the psychological level, implications of rigidity on the one hand, and of compliance and suggestibility on the other, would seem to point toward patient responses which are in the same direction as the neurophysiological level inhibitory potential which is also likely to be present.

Our clinical experience with psychiatric and medical clinic outpatients has also focused our attention upon such psychological aspects which seem to be of considerable relevance to an understanding of the side reactions reported by the two clinic groups in terms of the meaning that the drug treatment situation has for them (cf. 34). Medical clinic patients report, for example, both more frequent side reactions to and less improvement from imipramine or from placebo than is the case with patients from the psychiatric clinic. We believe that, because of the extreme passive-dependence of the

former patients, the only mode of expressing dissatisfaction with treatment to the all-powerful doctor, whose help and support they find essential, is through the reporting of troublesome symptoms whose onset occurred after the inception of treatment. Also, side reactions are probably more readily available to these patients as a means of expressing their dissatisfaction, since their somatic centering makes them more aware of slight alterations in bodily sensation (e.g., produced by imipramine) than is the case with psychiatric outpatients who are more interested in receiving help with their psychological problems. Also, patients who are hypersensitive to internal cues or signs may get frightened by somatic side reactions and therefore build up a psychological counter-reaction to the drug. By way of contrast Deprol, in low dosages, made medical patients sleep better and feel more relaxed; consequently, they felt that the doctor was helping them and so their high suggestibility furthered a "positive psychological attitude" toward drug and doctor. These patients returned and said how wonderful they felt.

Dropping out of therapy represents another reaction to drug treatment in which the effect of personality and attitudinal variables is of considerable importance. Patients with psychiatric problems who seek and accept drug treatment typically expect that a competent and all-knowing physician will prescribe for them a drug which will "make them feel better," that is, remove their symptoms. Hence, the fact that three times as many patients receiving placebo as patients receiving an active agent (meprobamate, Librium) terminate therapy after the first or second session is not surprising (42). However, it has also been found (42) that a larger proportion of primarily anxious than of primarily depressed patients drop out of therapy. The anxious patient is probably less tolerant of a period of no improvement and, as will be detailed in a subsequent paper, is more likely to give direct expression to his hostile feelings. In contrast, the self-directed hostility of the depressed patient is more likely to cause him to prolong a situation in which he experiences no improvement.

But the situation with regard to premature termination appears to be more complex. Freedman, *et al.* (17), report that patients evidencing less anxiety, hostility, and suspiciousness, and more energy and interpersonal interest proved to be the ones who more frequently dropped out of drug therapy. Similarly, Dymond (12), in characterizing patients dropping out of psychotherapy, notes that, in addition to possessing attributes relating to difficulties in giving verbal expression to feelings (a patient characteristic most likely predisposing to dropping out from the psychotherapy situation, cf. 22), these individuals see themselves as competitive, make demands upon themselves,



and try to cope with the world. Again, Imber, *et al.* (24) have found that greater suggestibility and membership in the middle socio-economic class as opposed to the lower socio-economic class are associated with remaining for a longer period in short-term psychotherapy. Thus it may be that a set of factors associated with dropping out has to do with the patient's discovering that drug therapy is not compatible with his needs for being independent and solving his problems through his own efforts. Hence, we expect those patients to be more likely to remain in drug treatment who fall within the center portion of the anxiety continuum, who are primarily oriented toward dependency gratification, and who give but limited overt expression to their hostility (44).

### C. THE TEST BATTERY

Hence, our experience in working with clinic populations, reinforced by our review of the literature, led us to feel that the development of a battery of tests for the measurement of personality and attitudinal factors might open the way to a program of research which would throw further light upon the interaction between patient characteristics and their mode of response to drug treatment. The considerations detailed in the foregoing discussion led us to select the following variables as those most likely to be related to the mode of drug response: anxiety, compliance, dependency, manifest hostility, self-acceptance, and attitude toward the doctor and toward parental figures. Before presenting a more formalized statement of the hypotheses to be tested in this research program, we shall describe the test battery assembled to measure these personality variables in our particular patient populations. Furthermore, we will report the outcome of a preliminary study carried out to verify first impressions about clinic differences, to test inter-relationships, among our variables and to provide evidence for the construct validity of the instruments we are employing.

Our patient population is composed of anxious, tense, irritable, worried and mildly depressed non-psychotic patients who come for help either to the Psychiatric Clinic or to the General Medical Clinic of the University of Pennsylvania. They are predominantly of the lower socio-economic class and frequently limited in intelligence, education, and ability to give verbal expression to their feelings.

While this population is an extremely important one in that it typifies the vast majority of individuals currently in need of psychiatric and pharmacological help, its limitations in intelligence, verbal ability, and education drastically curtail the types of measuring instruments which may be used to study the personality structure of its members. For example, the vocabulary

level of such questionnaire-type tests as the Minnesota Multiphasic Inventory and the Edwards Personal Preference Schedule is frequently too high. The small number of responses usually given to the Rorschach and the unembellished description of the cards which form the modal response to the TAT, mean that these techniques produce but sparse information and that quantitative indices of acceptable reliability cannot be derived from them.

We have attempted to select or devise a battery of tests which involve simple language, which require a direct and well delimited response from the patient, and which will provide us with dependable, quantitative indices of those personality factors which we believe important in predicting and understanding patient reactions to our drug treatment situation. In what follows, we shall describe each of the tests of our battery in its present form and detail our reasons for its selection.

*Taylor Anxiety Scale:* Since this instrument has been so widely used and extensively evaluated (4, 25, 48, 49), we do not feel that a review of this literature is necessary. Blake and Mouton (4) seem to do justice to the many studies with conflicting outcomes when they say, "Based on these several results, the outcome of whether or not the Taylor MAS constitutes an adequate psychometric measure of anxiety is a draw." Kern, Ewing, and Rickels (30) found significant Spearman rank order correlations ( $\rho = 0.66$ ) between the combined ratings of two psychiatrists and the Taylor MAS. Previous work with the instrument has convinced us that, even though it is not feasible to use the entire MMPI from which these items are selected, populations similar to the one taking part in the present study are sufficiently able to comprehend the language in which the items are couched to make the test, in this respect, appropriate for our subjects.

*Q-Sort:* We are using a slightly modified form of a procedure described by Hilden (20). Fifty items are to be sorted into five categories, once to describe the actual self, once to describe the ideal self. An unforced sorting distribution is used. The items are one of the fifty-item samples that Hilden has drawn at random from a population he has constructed by building item statements about all the adjectives at the sixth grade level or below, from the *Thorndike Century Senior Dictionary*, which can be used to describe people. Hilden (21) provides evidence for a high correspondence between actual-ideal self discrepancy measures based on 50 item samples, and similar measures based on his population of 1575 items. Thus, we are at once provided with a set of Q-sort items which sample a known universe and whose difficulty level is defined in such a way as to be appropriate for our population (cf. 10, 52).



There has been considerable discussion in the literature about the relative merits of forced as opposed to unforced sorting distributions (5, 9, 26, 36); and it was only after a careful weighing of the issues that we decided upon an unforced distribution. Those experimental results which seem to favor the use of a forced frequency distribution are based almost entirely upon the sortings of college student subjects, and are often concerned with the subject's description of political or other types of public figures (cf. 5, 36). We believe that tendency toward black-white dichotomies in regard to self and self-ideal are important characteristics of our patients which may be detected in the shapes of the sorting distributions they produce. This contention seems to gain some support from Jones (26), who found differences in distribution shape between college and mentally ill populations with the latter tending to give more extreme responses. We also felt that the added number of discriminations which might be made by those subjects who were willing or able to comply with a forced distribution procedure would not compensate for the total absence of data from the non-complying subjects of our intellectually limited and concretely oriented group.

The question of the contamination of Q-sort measures by social desirability measures represents another issue which has received much attention in the current literature (cf. 8, 10, 13, 14, 29, 31, 52). Edwards (13) and Kenny (29) have found substantial correlations between the social desirability of items and the frequency of their endorsement for both actual self and ideal self ratings (and this should be true in the latter case virtually by definition). However, Cowen, *et al.* (8), the only workers to the authors' knowledge who have correlated social desirability with actual self-ideal self discrepancy scores, found much lower correlations between such scores and social desirability than was the case when social desirability scores were correlated with actual or ideal self scores alone. We feel that such data provides some justification for our use of discrepancy score measures, without danger of extreme social desirability bias.

A final problem of considerable importance which deserves our attention concerns the role of defensiveness in the generation of low actual self-ideal self discrepancies. Low discrepancy scores may be a function of high self-satisfaction or a defensive refusal to face dissatisfaction. Studies involving subjects with only mild adjustmental problems (19, 50) report higher discrepancy scores in the poorly adjusted. But Block and Thomas (6), while in general finding greater maladjustment associated with greater discrepancy scores, note that those 85 subjects with the 10 lowest discrepancy scores obtain a higher score on the denial scale which has been found to be one

component of the HY scale of the MMPI. The comparisons made among the discrepancy scores of normals, neurotics, and psychotics by Friedman (18), Hillson and Worchel (23), and Zuckerman and Monashkin (54) would seem to suggest that defensive distortions are much more likely to occur with psychotic inpatient groups rather than neurotic outpatient groups. Since our patients are typically characterized by mild depressive or anxiety reactions, we believe that it is unlikely that defensiveness would take the form of enhancement of self-esteem and believe that lowered discrepancy scores might be taken as valid indicators of increments in self-acceptance.

*Hostility:* In order to measure the level of manifest or reported hostility, we have used a questionnaire technique developed by Buss and Durkee (7). We have employed 66 of the 75 items, deleting those which Buss and Durkee employed as a measure of guilt. Hostility items are divided into seven areas: assault, indirect hostility, irritability, negativism, resentment, suspiciousness, and verbal hostility. A factor analysis of the scale has yielded the two factors of attitudinal and verbal hostility, but considerable unique variance is left for the individual scales. We have selected this particular scale because the items are written in relatively simple language, item analysis techniques have been used in item selection, several devices have been employed to minimize the effect of social desirability, and opportunity is provided for using an overall as well as component hostility scores.

*Compliance:* A Rorschach technique modeled after that employed by McReynolds (37) and Keuthe (32) was employed for the measurement of compliance. Five possible responses have been selected for each of the 10 Rorschach cards.<sup>2</sup> Of these 50 responses 24 are F-plus, 24 are F-minus, and two are plus and minus, as given in the scoring norms provided by Beck (1). The subject is given Card I, a response is suggested and he is asked to say "yes" or "no" depending upon whether or not he feels that the suggested response can be seen in the blot before him. A response is then presented for each of the following nine cards in sequence; then the sequence of 10 cards is repeated four additional times until the 50 responses have been run through. The score is the total number of "yes" responses given by the subject. The instructions, intended to emphasize a compliance effect are as follows:

When people look at these cards they see a lot of different sorts of things on them. I'm going to read you a list of things a number of people have seen, and ask you whether you can see each one. I'll show you these cards one at a time and ask you a question about what each could be. For example, do you think that this could be a bat? (showing Card 1 and then proceeding with the rest of the cards).

<sup>2</sup> List of responses available on request.



Reliability coefficients of .82 and .70 have been reported by McReynolds and Kuethe respectively for this task. Since Kuethe found a significant correlation (.42) between the number of "yes" responses and productivity on a dull task done to please the experimenter and a correlation of only .02 when the same task was not done to please the experimenter, we feel that there is some evidence that a score on this test may represent a desire to please and to comply with the doctor on the part of our patients.

*Dependency:* Since the available tests of dependency seemed more appropriate for college student populations than for our patients, we decided to devise an instrument of our own which would require subjects to choose which statement of a pair was more like them. Twenty-five Item pairs made up of every possible combination of five statements descriptive of independent behavior with five statements descriptive of dependent behavior were used. These statements were obtained in the following manner: A group of 10 psychiatrists and psychologists were asked to list all the adjectives they could think of to be descriptive of dependent and independent behavior. When these lists were collected and tabulated, it was found that the great majority of adjectives so obtained would be well beyond the vocabulary level of our subjects. Inspection of the adjectives revealed that they could, with the exception of a group of independent items relative to active striving, be grouped under the three aspects of the dependency construct presented by Bernardin and Jessor (3): (a) Reliance on or the importance of approval from others; (b) reliance on others for help and assistance; and (c) a conformity to the opinions and demands of others. A number of phrases in simple language were written to encompass the content of the less readily comprehensible adjectives previously obtained. Five phrases descriptive of independent and five phrases descriptive of dependent behavior were then selected, the selection being based upon the ability of 39 psychologist and psychiatrist judges to agree that the items fell at the dependent or independent end of a dependency-independency scale. The same 39 judges also assigned the items to an 11 point Thurstone type scale with acceptable inter-judge agreement. Weighted and underweighted scoring systems will be studied to determine their relative merits as data from patient groups of sufficient size become available. The ten items are as follows:

*Independent*

- I feel sure of myself.
- I stand up for myself.
- I try hard to get ahead.
- I say what I think no matter what.
- I am likely to give orders to other people.

*Dependent*

- I look for help when a problem comes up.
- I give in easily.
- I get upset if somebody doesn't like me.
- I try to please others.
- I wait for good things to come to me instead of going after them.

The instructions for the 25 item paired comparison test are:

Read each pair of phrases and then put an X in front of the phrase that is *most like you*. For example, if you were six feet tall, you would mark this pair of phrases like this:

X am tall                      — am short

*Semantic Differential*: This instrument, developed by Osgood (39) and more fully reported on by Osgood, Tannenbaum and Suci (40), was employed to obtain a measure of attitude toward the doctor, attitude toward parental figures, and several aspects of attitude toward the self. A five-point scale, rather than the seven-point scale employed by Osgood, was selected as more appropriate to the intellectual level of our subject population. Since it was found by Downing, Moed and Wight (11) that such a five-point scale could be meaningfully employed by a group of children slightly below average in intellectual level and ranging in age from seven to 18 years, it was felt that such a score range could also be meaningfully used by the members of our patient population. Since it was also found (11) that the evaluative, potency, and activity factors of connotative meaning, which are the main dimensions of meaning content emerging from the test, vary from high to low in reliability in the order in which they are listed, the number of scales used to measure each factor was varied accordingly. Six scales were employed for the evaluative factor and three for the potency factor. Responses to activity factor scales were so inconsistent that it was decided to exclude data on this factor from our analysis.

The following scales were used to measure each factor: *Evaluative*—good-bad, friendly-unfriendly, wise-foolish, honest-dishonest, safe-dangerous, selfish-unselfish; *Potency*—strong-weak, hard-soft, brave-cowardly; *Activity*—fast-slow, excitable-calm. The concepts to be rated are:

MYSELF  
DOCTOR  
MY MOTHER  
MY FATHER  
MAN

WOMAN  
GOD  
GENTLEMAN  
CROOK  
BUM

GOD, GENTLEMAN, CROOK and BUM are to be used as "good" and "bad" stereotypes whose distance from MYSELF, DOCTOR, MY MOTHER, and MY FATHER may be measured in the semantic space of each individual (cf. 40). MAN and WOMAN are included to investigate the possibility that a comparison of MYSELF-MAN and MYSELF-WOMAN distances may be used as an indicator of the patient's sexual identification.



## D. SUMMARY

In this paper we have attempted to present the theoretical orientation toward the interaction of personality and attitudinal variables and response to drugs in psychiatric outpatients which has emerged from the integration of our clinical experience with findings reported in the literature. We have detailed our reasons for considering the personality variables of anxiety, compliance, dependency, manifest hostility, and attitude toward the doctor and parental figures as important factors to be considered in the understanding of such treatment responses as dropping out, side reactions, and placebo effects in the outpatient drug therapy situation. We have then described a battery of tests to be used in measuring these factors, and have presented the rationale which we have employed in their selection. In a subsequent paper we shall present evidence for the applicability of these tests in a hospital clinic setting, describe patterns of test inter-relationships which emerged in two clinic groups, and consider the implications for further research of our results.

## REFERENCES

1. BECK, S. J. Rorschach's Test, V 1, *Basic Processes*. New York: Grune & Stratton, 1950, pp. 227.
2. BEECHER, H. K. The powerful placebo. *J.A.M.A.*, 1955, **159**, 1602-1606.
3. BERNARDIN, A. C., & JESSOR, R. A construct validation of the EEPS with respect to dependency. *J. Consult. Psychol.*, 1957, **21**, 63-67.
4. BLAKE, R. R., & MOUTON, J. S. Personality. In Farnsworth, P. R. (Ed.), *Ann. Rev. Psychol.*, 1959, **10**, 203-232.
5. BLOCK, J. A comparison of the forced and unforced Q-sorting procedures. *Educ. Psychol. Measm.*, 1956, **16**, 481-493.
6. BLOCK, J., & THOMAS, H. Is satisfaction with self a measure of adjustment? *J. Abnorm. & Soc. Psychol.*, 1955, **51**, 254-259.
7. BUSS, A. H., & DURKEE, A. An inventory for assessing different kinds of hostility. *J. Consult. Psychol.*, 1957, **21**, 343-349.
8. COWEN, E. L., BUDIN, W., WOLITZKY, D. L., & STILLER, A. The social desirability of trait descriptive terms: A factor in the prediction of Q-sort. *J. Pers.*, 1960, **28**, 530-544.
9. CRONBACH, L. M., & GLESER, G. C. Assessing similarity between profiles. *Psychol. Bull.*, 1953, **50**, 456-473.
10. CROWNE, D., & STEPHENS, M. W. Self-acceptance and self-evaluative behavior: A critique of methodology. *Psychol. Bull.*, 1961, **58**, 104-121.
11. DOWNING, R. W., MOED, G., & WIGHT, B. W. Studies of disability: A technique for the measurement of psychological effects. *Child Devel.*, 1961, **32**, 561-575.
12. DYMOND, R. F. Adjustment changes in the absence of psychotherapy. *J. Consult. Psychol.*, 1955, **19**, 103-107.
13. EDWARDS, A. L. Social desirability and Q-sort. *J. Consult. Psychol.*, 1955, **19**, 462.
14. EDWARDS, A. L., & HORST, P. Social desirability as a variable in Q-technique studies. *Educ. Psychol. Measm.*, 1953, **13**, 620-625.

15. EYSENCK, H. J. Drug postulates, theoretical deductions, and methodological considerations. In Uhr, L., & Miller, J. G. (Eds.), *Drugs and Behavior*, New York: Wiley, 1960, pp. 452-464.
16. VON FELSINGER, J. M., LASAGNA, L., & BEECHER, H. K. Drug-induced mood changes in man: 2. Personality and reactions to drugs. *J.A.M.A.*, 1955, **157**, 1113-1119.
17. FREEDMAN, N., ENGELHARDT, D. M., HANKOFF, L. D., GLICK, B. S., KAYE, H., BUCHWALD, J., & STARK, P. Drop-out from outpatient psychiatric treatment. *A.M.A. Arch. Neur. & Psychiat.*, 1958, **80**, 657-666.
18. FRIEDMAN, I. Phenomenal, ideal, and projected conceptions of self. *J. Abnorm. & Soc. Psychol.*, 1955, **51**, 611-615.
19. HANLON, T. E., et al. Congruence of self and ideal self in relation to personality adjustment. *J. Consult. Psychol.*, 1954, **18**, 215-218.
20. HILDEN, A. H. Manual for Q-Sort and Random Sets of Personal Concepts. Webster Groves, Mo.: The author, 1954, pp. 19.
21. ———. Q-sort correlation: Stability and random choice of statements. *J. Consult. Psychol.*, 1958, **22**, 45-50.
22. HILER, E. W. The Sentence Completion Test as a predictor of continuation in psychotherapy. *J. Consult. Psychol.*, 1959, **23**, 544-549.
23. HILLSON, J. S., & WORCHEL, P. Self concept and defensive behavior in the maladjusted. *J. Consult. Psychol.*, 1957, **21**, 83-88.
24. IMBER, S. D., FRANK, J. D., GLIEDMAN, L. H., NASH, E. H., & STONE, A. R. Suggestibility, social class and the acceptance of psychotherapy. *J. Clin. Psychol.*, 1956, **12**, 341-344.
25. JENKINS, J. J., & LYKKEN, D. T. Individual differences. In Farnsworth, P. R. (Ed.), *Ann. Rev. Psychol.*, 1957, **8**, 79-112.
26. JONES, A. Distributions of traits in current Q-sort methodology. *J. Abnorm. & Soc. Psychol.*, 1956, **53**, 90-95.
27. JOYCE, C. R. B. The personality of healthy reactors and non-reactors to dummy treatment, in *Neuropsychopharmacology*. Elsevier Pub. Co., 1959, **1**, 448.
28. KAST, E. C., & LOESCH, J. Influence of the doctor-patient relationship on drug action. *Illinois Med. J.*, 1961, **119**, 390.
29. KENNY, D. T. The influence of social desirability on discrepancy measures between real self and ideal self. *J. Consult. Psychol.*, 1956, **20**, 315-318.
30. KERN, T. G., EWING, J. H., & RICKELS, K. The Hippuric Acid Index: Application of the Hippuric Acid Test as an index of human free anxiety. *J. Psychosom. Res.*, 1959, **3**, 211.
31. KOGAN, W. S., QUINN, R., AX, A. F., & RIPLEY, H. S. Some methodological problems in the quantification of clinical assessment by Q-array. *J. Consult. Psychol.*, 1957, **21**, 57-62.
32. KUETHE, J. L. The positive response set as related to task performance. *J. Pers.*, 1959, **27**, 87-94.
33. LASAGNA, L., MOSTELLER, F., VON FELSINGER, J. M., & BEECHER, H. K. A study of placebo response. *Am. J. Med.*, 1954, **16**, 770-779.
34. LINDEMANN, E., & CLARK, L. L. Modifications in ego structure and personality reactions under the influence of the effects of drugs. *Am. J. Psychiat.*, 1952, **56**, 561-567.
35. LINDEMANN, E., & VON FELSINGER, J. M. Drug effects and personality theory. *Psychopharmacologia*, 1961, **2**, 69.
36. LIVSON, N. H., & NICHOLS, T. F. Discrimination and reliability in Q-sort personality descriptions. *J. Abnorm. & Soc. Psychol.*, 1956, **52**, 159-165.
37. McREYNOLDS, P. Perception of Rorschach concepts as related to personality deviations. *J. Abnorm. & Soc. Psychol.*, 1951, **46**, 131-141.



38. MINER, J. B. Intelligence in the United States. New York: Springer, 1958.
39. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 197-237.
40. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Ill. Press, 1957, pp. 324.
41. RICKELS, K. Controlled outpatient studies with non-psychotic psychiatric patients. *Trans. 5th Ann. Res. Conf.*, V.A., 1960, **5**, 101-106.
42. RICKELS, K., BAUMM, C., & FALES, C. The influence of two different "sets" on placebo responses of neurotic outpatients. In preparation.
43. RICKELS, K. Placebo reactions and drop-outs in neurotic and depressed patients visiting psychopharmacology clinics. Unpublished manuscript.
44. RICKELS, K., CLARK, T. W., EWING, J. H., & KLINGENSMITH, W. C. Evaluation of tranquilizing drugs in medical outpatients. *J.A.M.A.*, 1959, **171**, 1649.
45. ROSENTHAL, D., & FRANK, J. D. Psychotherapy and the placebo effect. *Psychol. Bull.*, 1956, **53**, 294-302.
46. SHAGASS, C. Drug thresholds as indicators of personality and affect. In Uhr, L., & Miller, J. G. (Eds.), *Drugs and Behavior*, Wiley, 1960. pp. 399-404.
47. SHAGASS, C., & LIPOWSKI, Z. J. Effect of methedrine on critical flicker fusion and its relation to personality and affect. *J. Nerv. Ment. Dis.*, 1958, **127**, 407-416.
48. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abnorm. & Soc. Psychol.*, 1953, **48**, 285-290.
49. ———. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303-320.
50. TURNER, R. H., & VANDERLIPPE, R. H. Self-ideal congruence as an index of adjustment. *J. Abnorm. & Soc. Psychol.*, 1958, **57**, 202-206.
51. WIKLER, A. Mechanisms of action of drugs that modify personality function. *Am. J. Psychiat.*, 1952, **56**, 590-599.
52. WITTENBORN, J. R. Contributions and current status of Q-methodology. *Psychol. Bull.*, 1961, **58**, 132-142.
53. WOLF, S., & PINSKY, R. H. Effects of placebo administration and occurrence of toxic reactions. *J.A.M.A.*, 1954, **155**, 339-341.
54. ZUCKERMANN, M., & MONASHKIN, I. Self-acceptance and psychopathology. *J. Consult. Psychol.*, 1957, **21**, 145-148.

University of Pennsylvania  
School of Medicine  
Department of Psychiatry  
Philadelphia 4, Pennsylvania

# PERSONALITY AND ATTITUDINAL CORRELATES OF RESPONSE TO DRUG TREATMENT IN PSYCHIATRIC OUTPATIENTS: II. PRELIMINARY INVESTIGA- TION AND IMPLICATIONS FOR FURTHER RESEARCH\*<sup>1</sup>

*University of Pennsylvania School of Medicine*

---

R. W. DOWNING AND K. RICKELS

---

## A. INTRODUCTION

In a previous paper (Rickels and Downing, 9), we presented our conceptual orientation of the role of personality and attitudinal factors in the response to treatment of neurotic outpatients with tranquilizing drugs. In the present paper we shall describe the results of a preliminary investigation involving small groups of neurotic patients from a Psychiatric Outpatient Clinic and from a General Medical Clinic. We have obtained further data from a sample of fifteen hospitalized schizophrenics. Comparisons and contrasts of this group with the other groups will be introduced at various points throughout the paper wherever it is considered appropriate.

The aims of the present preliminary study are threefold. Firstly, to determine the applicability of the test battery described in our previous paper (Rickels and Downing, 9) to the lower socio-economic class and lower intelligence group, of which our population consists; secondly, to study the inter-relations among the several tests, so as to gather information concerning the construct validity (cf. Chrombach and Meehl, 1955) of our measuring instruments; and finally, to ascertain what evidence there might be that our two clinic groups differ sufficiently in personality structure to make it likely that they might respond differentially in drug treatment.

As we have previously stated, Psychiatric Clinic patients are somewhat oriented toward considering their problems as psychological in origin and are more likely to expect psychologically focused treatment. On the other hand, Medical Clinic patients tend to be more somatically focused and expect to receive medication as their sole treatment.

---

\* Received in the Editorial Office on June 18, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Work supported by USPHS Grant MY 2934.



## B. METHOD

We shall provide here only a brief description of the tests that we have employed, since these tests have previously been discussed in considerable detail (9). The test battery is intended to provide measures of manifest anxiety, compliance, dependency, manifest hostility, as well as attitudes toward the self and toward the doctor.

*Taylor Manifest Anxiety Scale*—50 items descriptive of psychological and physiological manifestations of anxiety which were selected by Taylor (10) from the MMPI.

*Compliance Rorschach*—A technique in which five responses from each of the 10 Rorschach cards are suggested to the patient, one at a time. The patient is required to respond "Yes" or "No" to indicate whether he feels that each percept is an appropriate response to a given Rorschach card. The number of "Yes" responses in this test has been found by Kueth (5) to be related to level of performance on a monotonous task when the subject's motivation was solely to please the experimenter.

*D-I Inventory*—A 25-item paired comparison test, devised by ourselves, in which each of five items agreed by a group of judges to be descriptive of dependent behavior is paired with each of five items similarly agreed to be descriptive of independent behavior. The subject's task is to decide which item in each of the 25 pairs is most like him.

*H-S Inventory*—A set of 70 True-False questions taken from an inventory constructed by Buss and Durkee (1). In addition to a total hostility score, measures of the following seven specific types of hostility may be obtained: Assault, Indirect Hostility, Irritability, Negativism, Resentment, Suspiciousness, and Verbal Hostility. The test has also been factor analyzed to yield factors of Attitudinal Hostility, and Motor Hostility.

*Q-Sort*—This consists of 50 statements, sampling the personality domain, which are to be sorted on a 5-point scale so as to describe the actual and the ideal self. The sampling procedure and the items used were devised by Hilden (4).

*Semantic Differential*—Employing the technique devised by Osgood (7), we have selected 10 concepts which are rated along a 5-point scale for six adjective pairs representing the Evaluative dimension, and three pairs representing the Potency dimension of connotative meaning. The concepts chosen are MYSELF, DOCTOR, MY MOTHER, MY FATHER, GENTLEMAN, GOD, BUM, CROOK, MAN, and WOMAN. In addition to obtaining direct measures of the connotative meaning of such

terms as MYSELF and DOCTOR, it is possible to measure the similarity of meaning between such concepts, on the one hand, and positive (GENTLEMAN, GOD) and negative (BUM, CROOK) stereotypes, on the other.

The patient population consists of urban lower socio-economic class individuals who have come for treatment either to the Psychiatric Clinic or to the Medical Clinic of a large university hospital. They are psychoneurotics and character neurotics.

TABLE 1  
AGE, RACE AND SEX FOR PSYCHIATRIC CLINIC, MEDICAL CLINIC AND  
HOSPITALIZED SCHIZOPHRENIC PATIENTS

	Race		Sex		Mean age
	White	Negro	Male	Female	
Psychiatric clinic (n = 13)	12	1	2	11	36.3
Medical clinic (n = 13)	4	9	2	11	50.8
Hospitalized schizophrenics (n = 15)	9	6	8	7	38.5

Data concerning the sex, age, and race of the patients in our sample are shown in Table 1. The preponderance of females over males reflects the sex distribution of the population concerned. Since the clinic groups obviously differ in age, it is important to establish that clinic differences in test scores are not merely a function of age differences. With this view in mind, correlations were obtained between age and test score for each of the tests in the battery. These correlations are presented in Table 2. It will be noted that none of them attains statistical significance, and that most of them are of negligible size. Thus, it seems safe to infer that clinic differences are not merely a function of age.

TABLE 2  
RANK ORDER CORRELATIONS ( $Rho$ ) OF TEST SCORES WITH AGE AND WITH VOCABULARY  
LEVEL FOR PSYCHIATRIC AND MEDICAL CLINICS COMBINED

		Age	$p$	Vocabulary	$p$
Taylor MAS	(n = 25)	-.03	—	-.04	—
Rorschach Compliance	(n = 24)	.22	—	.01	—
D-I inventory	(n = 22)	.04	—	.04	—
H-S inventory	(n = 25)	-.04	—	-.19	—
Q-sort Self discrepancy	(n = 23)	-.08	—	-.06	—



It will also be noted that the clinics differ markedly in their racial composition. Psychiatric Clinic patients are predominantly a white group, while about two thirds of the patients in our Medical Clinic sample are Negro. While the sample sizes of the groups in this preliminary investigation are, of course, too small to provide a clear cut differentiation between the influences of race and that of clinic population, we feel that some indication of the effects of these two factors may be had by studying separately the patterns of test scores and their inter-relationships in clinic groups and in racial groups. With this view in mind, we shall tabulate our data both by clinic and by race, and we shall provide a more systematic study of the effects of clinic and racial factors in our major investigation.

The difference in intelligence between our two clinic groups as measured by the Thorndike Vocabulary Test (6), is of borderline significance. The mean Vocabulary Test score for the Psychiatric Clinic is 11.70; that for the Medical Clinic is 10.97. The  $t$ -test value for this difference is 1.41 with a significance value between the 10 and 20 per cent levels for 27 degrees of freedom.<sup>2</sup>

It was felt important to establish that intellectual level does not represent a major component of variance in the test scores. Therefore, correlations between the score on each test and the Thorndike Vocabulary Test score are also given in Table 2. It will be noted that none of these correlations reaches statistical significance.

## C. RESULTS AND DISCUSSION

### 1. *Attitudes toward Self*

Tables 3 through 6 present the mean test scores by clinic and by race and give the  $t$  and  $p$  values associated with mean differences.

<sup>2</sup> Because of practical difficulties in scheduling, and because of limitations in intelligence, all of the patients for whom testing was attempted could not be given a complete test battery. Intellectual limitation represents a negligible factor in restricting the test ability of Psychiatric Clinic patients, while about one third of our Medical Clinic patients are not completely testable. Adding the Vocabulary Test scores of a group of Medical Clinic patients unable to take our full battery to the scores of testable patients, we find that the obtained sample of 22 cases from the Medical Clinic produces a mean Vocabulary Test score of 8.64; as compared with a mean score of 11.29 for a sample of 17 cases from the Psychiatric Clinic ( $t = 2.43, p < .02$ ).

However, we do not feel that this limitation in sampling interferes with our aim of demonstrating those forms of personality differences among the two clinic populations which would be reflected in differences in response to drugs between the two groups. It seems to us very likely that those personality characteristics found among the testable sample of Medical Clinic patients exist in an even more exaggerated form in those Medical Clinic patients too intellectually limited for testing.

Differences in scores on the Q-Sort self-discrepancy measure (Table 3) and on the Semantic Differential evaluative factor score for MYSELF (Table 5), both indicate that Psychiatric Clinic patients are much less accepting of themselves than is the case for Medical Clinic patients. It will be noted by consulting Tables 4 and 6 respectively, that white and Negro groups differ significantly on neither of these values. Thus, we may conclude that the difference in self-acceptance between clinics is not merely a reflection of differences in the racial composition of clinic groups. Furthermore, the significantly lower ratings of MYSELF on the potency factor made by the Negro group would seem to suggest that this group, composed predominantly of lower socio-economic class females are able to internalize a view of themselves as somewhat lacking in power without consciously lowering the value they place upon themselves.

TABLE 3  
MEAN SCORES ON TESTS FOR PSYCHIATRIC AND MEDICAL CLINICS

Test	Psychiatric Clinic	Medical Clinic	t	df	P
Taylor MAS	25.8	18.6	1.99	23	.10 > <i>p</i> > .05
Rorschach Compliance	9.9	12.7	1.48	23	.20 > <i>p</i> > .10
D-I inventory	10.0	2.5	.91	21	—
Hostility inventory	43.3	37.3	.92	24	—
Q-sort self-discrepancy	60.2	37.0	3.48	22	.01 > <i>p</i> > .001

TABLE 4  
MEAN SCORES ON TESTS FOR WHITE AND NEGRO GROUPS

Test	Mean		t	df	P
	White	Negro			
Taylor MAS	22.1	19.4	.54	23	—
Rorschach Compliance	11.2	11.6	.14	23	—
D-I inventory	8.3	2.4	.71	21	—
Hostility inventory	37.1	45.4	1.16	24	—
Q-sort self-discrepancy	50.8	43.2	.91	22	—

When the mean self-discrepancy scores from the Q-Sort and the mean self-evaluation scores from the Semantic Differential obtained by the Psychiatric and Medical Clinic patients are compared with the same scores from a group of hospitalized schizophrenics, we discover some possible differences in self orientation among the three groups. The schizophrenics obtain a mean self-discrepancy score on the Q-Sort of 60.5. This figure does



not differ from the mean obtained by the Psychiatric Clinic patients (60.2), but it differs ( $.02 > p > .01$ ) from the mean value of 37.0 obtained by Medical Clinic patients. Thus, it appears that our anxious and depressed Psychiatric Clinic outpatients and our hospitalized schizophrenics both see a large difference between themselves as they actually are and themselves as they would like to be, whereas Medical Clinic patients report a much smaller difference between actual and ideal selves.

However, when schizophrenics are requested to make an out-right evaluation of the "goodness" of themselves (as is requested on the Semantic Differential) it now appears that they share the same defensive pattern, perhaps denial, with the Medical Clinic patients. The schizophrenic self-evaluation mean score of .94 does not differ significantly from the mean score of Medical Clinic patients of 1.25, while it does show a borderline significance level in its difference from the Psychiatric Clinic mean of .55 ( $t = 1.62$ ,  $.20 > p > .10$ ).

It is tempting to speculate that Medical Clinic patients share with schizophrenics a propensity toward the use of denial as a defense, while the Psychiatric Clinic patients share with schizophrenics a marked tendency to make demands upon the self, with attendant guilt and anxiety. The finding by Hankoff *et al.* (3) that differences in hospitalization rate between outpatient schizophrenics treated with tranquilizing drugs and those treated with placebo was greater for "deniers" than for "acceptors" of mental illness may have some relevance to our situation. It suggests to us the tentative hypothesis that placebo-drug differences in improvement rates might be greater for the denial-prone Medical Clinic patients than it is for the more or less acceptance-prone patients of the Psychiatric Clinic.

## 2. *Anxiety, Compliance, Dependency and Hostility*

Although Psychiatric Clinic patients obtain a higher mean anxiety score than the Medical Clinic patients, within group variation is sufficiently great that this difference does not reach significance at the five per cent level when measured by the *t*-test. However, when Psychiatric and Medical Clinic patients are classified as being either above or below the median anxiety score for the two groups combined, nine Psychiatric Clinic patients are found to be above this combined median, and only three below. Only three Medical Clinic patients are above the combined median, and nine below (see Table 7, p. 353). Fisher's exact test based upon these figures reveals that the two distributions differ at the two per cent level.

TABLE 5  
MEAN SCORES FOR SEMANTIC DIFFERENTIAL CONCEPTS RATED ON EVALUATIVE (E) AND POTENCY (P) FACTORS  
FOR PSYCHIATRIC AND MEDICAL CLINICS

Concept		Psychiatric clinic	Medical clinic	t	df	p
MYSELF	(n = 13)		(n = 15)			$p < .001$
E		.55	1.25	3.50	26	—
P		— .15	— .02	.54	26	—
DOCTOR	(n = 13)		(n = 15)			$p < .02$
E		1.37	1.80	2.53	26	$.20 > p > .10$
P		.74	1.04	1.50	26	—
MOTHER	(n = 9)		(n = 13)			—
E		1.44	1.50	.25	20	—
P		.19	.31	.46	20	—
FATHER	(n = 9)		(n = 13)			—
E		1.26	1.49	.64	20	—
P		.72	.54	.61	20	—
GENTLEMAN	(n = 9)		(n = 13)			—
E		1.24	1.04	.61	20	—
P		.67	.44	1.06	20	—
GOD	(n = 8)		(n = 12)			—
E		1.83	1.90	.50	18	—
P		1.00	1.00	.00	18	—
BUM	(n = 9)		(n = 13)			—
E		— .13	— .41	.85	20	$.20 > p > .10$
P		— .52	— .98	1.47	20	—
CROOK	(n = 13)		(n = 15)			—
E		1.49	1.48	.04	26	$.10 > p > .05$
P		.26	.33	1.97	26	—
MAN	(n = 13)		(n = 15)			$.20 > p > .10$
E		.36	.73	1.42	26	$p < .05$
P		.08	.73	2.17	26	—
WOMAN	(n = 13)		(n = 15)			$.20 > p > .10$
E		.37	.74	1.32	26	—
P		.05	.11	.35	26	—



TABLE 6  
MEAN SCORES FOR SEMANTIC DIFFERENTIAL CONCEPTS RATED ON EVALUATIVE (E) AND POTENCY (P)  
FACTORS FOR WHITE AND NEGRO GROUPS

Concept		White	Patients	Negro	t	df	p
MYSELF	(n = 15)		(n = 12)				
E		.84		.89	.19	25	
P		.08		— .30	2.70	25	.02 > p > .01
DOCTOR	(n = 15)		(n = 12)				
E		1.66		1.55	1.00	22	
P		.80		.97	.89	22	
MOTHER	(n = 12)		(n = 10)				
E		1.58		1.35	1.00	20	
P		.36		.13	1.28	20	
FATHER	(n = 12)		(n = 10)				
E		1.65		1.08	.51	20	
P		.64		.67	.12	20	
GENTLEMAN	(n = 12)		(n = 10)				
E		1.14		1.10	.18	20	
P		.67		.44	.96	20	
GOD	(n = 10)		(n = 10)				
E		1.97		1.78	1.60	18	.20 > p > .10
P		.73		1.10	1.68	18	.20 > p > .10
BUM	(n = 12)		(n = 10)				
E		— .32		— .30	.02	20	
P		—1.00		— .73	.84	20	
CROOK	(n = 15)		(n = 12)				
E		—1.65		—1.27	1.42	25	.20 > p > .10
P		— .28		.16	1.67	25	.20 > p > .10
MAN	(n = 15)		(n = 12)				
E		.57		.57	.00	25	
P		.42		.66	.81	25	
WOMAN	(n = 15)		(n = 12)				
E		.72		.33	1.65	25	.20 > p > .10
P		.17		— .05	1.21	25	

The extreme variability responsible for the lack of significant difference between means is due to the fact that the scores of the three Medical Clinic patients which fall above the combined median are close to the upper end of the anxiety score distribution. Thus, while three fourths of the Medical Clinic patients (a significant proportion) obtain low anxiety scores, the three with high scores are apparently extremely anxious individuals, and probably represent people with a defensive organization differing from that of the rest of the group.

TABLE 7  
NUMBER OF PATIENTS FALLING ABOVE AND BELOW THE COMBINED GROUP MEDIAN  
FOR ANXIETY IN THE MEDICAL AND PSYCHIATRIC CLINICS

	Psychiatric Clinic	Medical Clinic
Above median	9	3
Below median	3	9

Neither the *t*-test nor the chi-square test described above reveal a significant difference in anxiety level between white and Negro groups. Thus, once more the difference between clinics does not seem to be merely a function of racial differences in anxiety level.

In consonance with our previous speculations (cf. p. 350), the mean anxiety score for our schizophrenic group is 23.8, a figure closer to the Psychiatric Clinic mean of 25.8, than to the Medical Clinic mean of 18.6. However, the mean anxiety score for the schizophrenic group does not differ significantly from that of either the Psychiatric or the Medical Clinic.

While Medical Clinic patients attain a higher mean dependency and compliance score and the Psychiatric Clinic patients attain a higher mean hostility score, none of these differences reaches statistical significance. Comparison of the data in Tables 3 and 4 reveals that all three of these differences represent less extreme contrasts between white and Negro groups than they do between Psychiatric and Medical Clinic groups. This leads us to expect that, should larger samples enable us to detect differences in regard to these variables between clinics, they would occur above and beyond differences in the racial compositions of the patients who make up the clinic populations.

### 3. Doctor Attitude and Other Factors From the Semantic Differential

Inspection of Table 5 reveals some interesting inter-clinic differences in regard to attitude toward the DOCTOR. Both groups place the DOCTOR on the "Good" side of the "Good-Bad" continuum of the Semantic Differen-



tial evaluative scale, but the estimate of the "goodness" of the DOCTOR is significantly higher for the Medical than for the Psychiatric Clinic group. Similarly, Medical Clinic patients tend to view the DOCTOR as more "potent," but this difference is of only borderline significance.

A reflection of a difference in life orientation between Psychiatric and Medical Clinic patients may possibly be seen in the fact that Medical Clinic patients attribute a higher evaluative score to seven of the ten concepts of the Semantic Differential ( $p = .11$ , sign test). A similar contrast does not exist with regard to the potency factor: Psychiatric Clinic patients rate four concepts as more potent, Medical Clinic patients rate five as more potent and there is one tie. It might be speculated that the defensive orientation of the Medical Clinic patient allows him consciously to regard the people of his social world as "better," i.e., to feel more positive toward them, than is the case with the Psychiatric Clinic patient.

When the Negro and white groups are compared with reference to their performance on the Semantic Differential (Table 6), it is found that only one of the twenty group differences reaches an acceptable level of significance; the Negro group rates MYSELF as less potent than does the white group. It should be noted that the group means for the evaluative factor rating of MYSELF are practically identical, as was stated previously (p. 349). This may reflect the fact that this group of Negroes (predominantly female) is able to assimilate the status of weakness assigned to them by the white majority, without a marked devaluation of the self.

#### 4. *Test Inter-relationships*

Inter-relationships among the tests within the two clinic and the two racial groups are presented in Table 8 and Table 9 respectively. It will be noted that within both clinic groups and within the white group, scores of the Taylor MAS, the Hostility Inventory, and the Q-Sort all tend to be positively inter-correlated. With the Negro group two of these three correlations are statistically significant, with that between Q-Sort and Taylor anxiety score just failing to reach significance at the five per cent level.

Thus, the patient with higher manifest anxiety tends to report himself as being more hostile and tends to have a greater discrepancy between his actual and his ideal self. This pattern is partially replicated with our group of fifteen schizophrenics, in that, a correlation of .42 ( $p < .05$ ) is obtained between Q-Sort self-discrepancy score and total hostility score, and a correlation of .63 ( $p < .01$ ) is obtained between Taylor manifest anxiety score and Q-Sort self-discrepancy score. In this group, the correlation

TABLE 8  
CORRELATIONS BETWEEN TESTS FOR PSYCHIATRIC AND MEDICAL CLINICS<sup>1</sup>

	2	3	4	5
<i>Psychiatric Clinic</i>				
1. Taylor MAS	11/—13	12/.52*	12/.65**	10/.67**
2. Rorschach Compliance		10/—01	12/—06	11/.19
3. D-I inventory			10/.37	10/.54*
4. Hostility inventory				11/.65**
5. Q-sort self-discrepancy				e
<i>Medical Clinic</i>				
1. Taylor MAS	13/.21	13/.26	13/.67**	13/.46*
2. Rorschach Compliance		13/.58*	13/—04	13/—04
3. D-I inventory			13/—21	13/—05
4. Hostility inventory				13/.67**
5. Q-sort self-discrepancy				

<sup>1</sup> Numbers above the diagonal indicate the number of cases, numbers below the diagonal, the values of Spearman rank order correlation coefficients.

\* Significant at less than the 5 per cent level.

\*\* Significant at less than the 1 per cent level.

TABLE 9  
CORRELATIONS BETWEEN TESTS FOR NEGRO AND WHITE GROUPS<sup>1</sup>

	2	3	4	5
<i>Negro Group</i>				
1. Taylor MAS	10/.32	10/.37	10/.74*	10/.52
2. Rorschach Compliance		10/.47	10/—03	10/—09
3. D-I inventory			10/.03	10/—08
4. Hostility inventory				10/.78**
5. Q-sort self-discrepancy				
<i>White Group</i>				
1. Taylor MAS	14/—37	12/.31	15/.78**	13/.69**
2. Rorschach Compliance		13/.04	15/—22	14/—08
3. D-I inventory			13/.25	13/.31
4. Hostility inventory				14/.81**
5. Q-sort self-discrepancy				

<sup>1</sup> Numbers above the diagonal indicate the number of cases, numbers below the diagonal, the values of Spearman rank order correlation coefficients.

\* Significant at less than the 5 per cent level.

\*\* Significant at less than the 1 per cent level.



between hostility and anxiety scores is, however, only .32, and fails to reach statistical significance. When the data for all three clinic groups are pooled, the three pairs of correlations among the three tests all reach significance at the one per cent level.

This set of findings seems in general in accord with those of the Rogerian group (cf. 11), that a tendency to dislike others is accompanied by a tendency to dislike the self and to fail to attain self-actualization, with attendant feelings of anxiety and dysphoria. As a larger amount of data becomes available, we think that it would be helpful in understanding this repeated pattern of correlation to item analyze the material of the Q-Sort against scores on anxiety and hostility scales.

Let us now consider a correlation pattern, which reveals a marked contrast between our Psychiatric and Medical Clinic groups (see Figure 1). In the Psychiatric Clinic, dependency score correlates both with anxiety level and with the actual-ideal self-discrepancy score derived from the Q-Sort; *i.e.*, the more dependent the Psychiatric Clinic patient reports himself to be, the higher is his anxiety score likely to be and the greater the dissatisfaction with self that is likely to be reflected in his Q-Sort. By way of contrast, dependency scores for Medical Clinic patients show a zero order correlation (.05) with Q-Sort performance, and only a moderate and non-significant correlation (.26) with anxiety level. Furthermore, within this group, there exists a correlation of .58 ( $p < .05$ ) between dependency and compliance measures. The correlation co-efficient providing an index of this relationship for the Psychiatric Clinic patients is of zero order ( $-.01$ ).

Our tentative interpretation of these findings is as follows: Psychiatric Clinic patients are unable to accept their dependency feelings, and hence deal with them in such a way that no systematic relationship emerges between them and the extent to which they behave in a compliant manner. On the other hand, Medical Clinic patients, perhaps as a result of their emphasis on somatizing defenses, are able to accept their dependent feelings and consequently express them more directly in compliant behavior. This interpretation seems to gain support from the fact that the relationship between dependency and hostility is negative ( $\rho = -.21$ ) for Medical Clinic patients, and positive ( $\rho = .37$ ) for Psychiatric Clinic patients. Although neither correlation co-efficient reaches statistical significance, this difference in direction suggests a trend for greater hostility to accompany greater dependency in Psychiatric Clinic patients and for less hostility to appear with greater dependency in Medical Clinic patients.

At the present time our data do not allow us to evaluate the relative

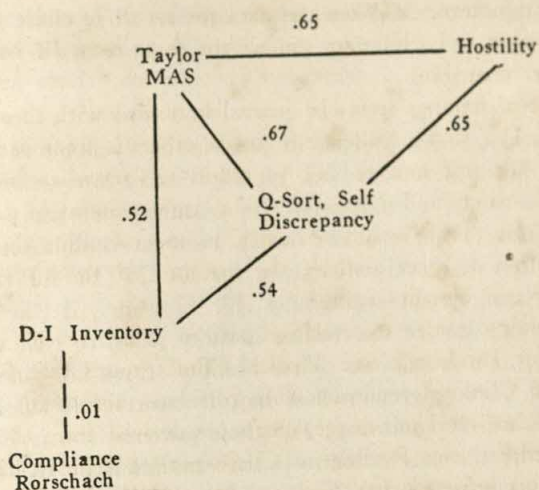
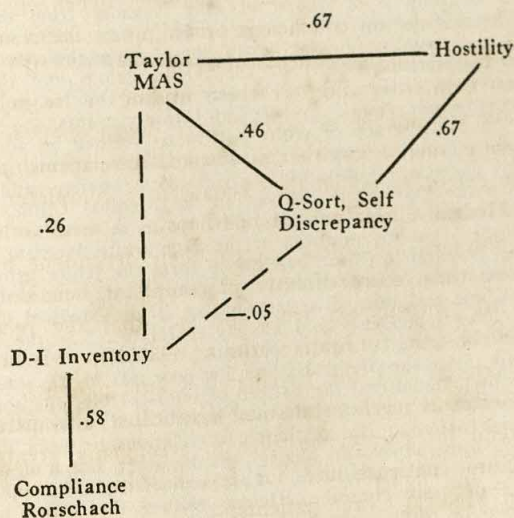
*Psychiatric Clinic**Medical Clinic*

FIGURE 1  
 DIAGRAMMATIC REPRESENTATION OF RELATIONSHIPS AMONG TEST VARIABLES  
 IN THE MEDICAL AND PSYCHIATRIC CLINICS



importance of racial differences as factors determining these differences in correlational pattern. However, there are several considerations which lead us to feel that the differences in defensive orientation of the two clinic groups may be making a significant contribution, above and beyond racial factors, to the patterns described immediately above.<sup>3</sup>

Although a sufficient number of cases is not available in this preliminary study to permit a more systematic investigation of interactional effects among the variables under consideration, an example may serve to demonstrate their presence. Thus, the correlation between compliance score and the semantic differential evaluative score for DOCTOR for the Medical and Psychiatric Clinics combined is only .15. However, if this combined group of patients is divided at the median hostility score, so that we consider separately one group of patients whose hostility scores fall above the combined median, and another group whose hostility scores fall below the combined median, we find the following: For those patients who report themselves to be less hostile, the correlation between evaluation of the DOCTOR and compliance score is .69 ( $.02 > p > .01$ ). However, for the more hostile patients, the corresponding correlation is only .02.

It thus appears that individuals describing themselves as less hostile give expression to their compliance through a high evaluation of the DOCTOR,

---

<sup>3</sup> Our Negro subgroup is composed of nine Negroes from the Medical Clinic and only one Negro from the Psychiatric Clinic. Thus, the results for the Negro subgroup must be virtually identical with those from the Medical Clinic. The pattern of inter-correlations shared by the Negro group and the Medical Clinic group may result from either factors characteristic of the Negro sub-culture or the pattern of psychodynamics found in Medical Clinic patients.

On the other hand, the group of white patients is formed by adding four Medical Clinic patients to a group of twelve Psychiatric Clinic patients. It will be recalled that dependency correlates substantially with both anxiety and Q-Sort self-discrepancy score in the Psychiatric but not in the Medical Clinic. If the defensive pattern of the Medical Clinic is found in its white as well as its Negro patients, it might be expected that the addition of the four white Medical Clinic patients to the twelve white Psychiatric Clinic patients to form the white group would produce a pattern of correlations which reflected a movement away from the pattern seen in the Psychiatric Clinic and toward a pattern seen in the Medical Clinic. This seems to be the case. Thus dependency and anxiety correlate .52 in the Psychiatric Clinic, but only .31 for the group of white patients. Similarly, dependency and self-discrepancy correlate .54 in the Medical Clinic but only .31 for the white group.

Finally, it has been found that the pattern of inter-correlations among the variables here under discussion is the same for our hospitalized schizophrenics as it is for the Medical Clinic patients; i.e., negligible correlations between dependency on the one hand and self-discrepancy and anxiety, on the other, but a significant correlation between dependency and compliance. This schizophrenic group is composed of nine whites and only six Negroes. Hence, neither the white patients from the Medical Clinic nor the white schizophrenics behave as do the white Psychiatric Clinic patients.

while patients describing themselves as more hostile do not. Compliance, as measured here, seems to have different implications for behavior depending upon the concomitant level of other variables.

#### D. COMMENTS: IMPLICATIONS FOR FURTHER RESEARCH

Further work now in progress will provide the opportunity for cross-validating our findings concerning clinic differences in personality factors and for relating these differences to factors in treatment response. We anticipate that the dependent, illness-denying Medical Clinic group, which is more strongly oriented toward the doctor as a good and powerful figure, will respond positively to pharmacological agents which alleviate primarily insomnia and other somatic symptoms, and will be most influenced by a suggestive therapeutic role. We expect that since these patients accept their dependent orientation they will not be troubled by side effects involving sleepiness and loss of energy. However, because of their somatic focusing, we anticipate that they will respond disproportionately to side effects themselves somatic in nature (e.g., dizziness, headaches, nausea, dry mouth). In earlier double blind drug evaluations we found that Medical Clinic patients complain less about drowsiness with Deprol and more about somatic side reactions with imipramine than Psychiatric Clinic patients (8).

On the other hand, we expect that Psychiatric Clinic patients who are more anxious, more worried about dependency and less self-accepting, will respond most favorably to doctor-patient interaction (psychotherapy), combined with a tranquilizing agent or even placebo, as long as this agent does not produce side reactions disturbing to them (e.g., sleepiness, drowsiness, and loss of energy). We expect, however, that somatic side effects will be less threatening to them, as these symptoms have less dependent meaning to the patient, than, for example, has a symptom such as drowsiness.

We also are prepared to test a number of hypotheses relating response to treatment to personality and attitudinal factors irrespective of the clinic membership of the patients in whom such factors manifest themselves. Thus, we expect that patients with excessive side reactions to either drug or placebo will prove themselves to be hostile, but dependent, individuals who express positive attitudes toward the doctor but can express their resentment only in such indirect ways as by complaining about side effects of the medication or by dropping out. Again, we expect premature termination of therapy to be found more frequently in patients either too dependent to obtain dependency satisfaction or too independent to need it. Finally, we expect symptomatic improvement in patients who are initially sufficiently free from



such pathological extremes on the variables measured that they can find some satisfaction in the treatment situation.

#### E. SUMMARY AND CONCLUSIONS

A test battery designed to provide measures of manifest anxiety, compliance, dependency, manifest hostility, and attitudes toward the self and the doctor, was administered to a group of 13 Psychiatric Clinic patients, 13 Medical Clinic patients, and 15 hospitalized schizophrenics. The application of the test battery proved feasible within the context of the clinic treatment situation. The test could be meaningfully responded to by a sufficiently large portion of the lower socio-economic class patients composing our treatment groups to make meaningful a more extensive and systematic investigation involving these instruments.

In our preliminary study, Psychiatric Clinic patients were found to be more anxious and less self-accepting than patients of the Medical Clinic, and there was a trend toward greater compliance in the Medical Clinic group. While the group did not differ in reported dependency, there was considerable evidence that this variable played a different role in patients from the two clinics. For Psychiatric Clinic patients, greater dependency was associated with higher anxiety and with lower self-acceptance, and there was a trend for it to be related positively with the quantity of reported hostility. Also, for these Psychiatric Clinic patients, dependency was not systematically related to our measure of compliance behavior.

For Medical Clinic patients, dependency was not related to self-acceptance. It tended to be negatively related to reported hostility, and was positively related to our measure of compliance behavior. There is some evidence that these different patterns of correlation result from differing defensive organizations of the patients in the two groups, rather than from differences in age and racial factors. The data would seem to indicate that dependency is directly expressed through compliant behavior when dependency is an acceptable element in one's self-concept, but that their direct expression is not present when one does not accept himself as a dependent individual.

Preliminary investigation indicates that the significance of the compliance score depends upon the concomitant level of other variables such as hostility and anxiety. We have discussed the implications of our findings for further research.

#### REFERENCES

1. Buss, A. H., & Durkee, A. An inventory for assessing different kinds of hostility. *J. Consult. Psychol.*, 1957, **21**, 343-349.

2. CHROMBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, **52**, 281-302.
3. HANKOFF, L. D., ENGELHARDT, D. M., FREEDMAN, N., MANN, D., & MARCOLIN, R. Denial of illness in schizophrenic outpatients: Effects of psychopharmacological treatment. *Arch. Gen. Psychiat.*, 1960, **3**, 657.
4. HILDEN, A. H. Q-sort correlations: Stability and random choice of statements. *J. Consult. Psychol.*, 1958, **22**, 45-50.
5. KUETHE, J. L. The positive response set as related to task performance. *J. Pers.*, 1959, **27**, 87-94.
6. MINER, J. B. Intelligence in the United States. New York: Springer, 1953.
7. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Ill. Press, 1961, pp. 370.
8. RICKELS, K. Controlled outpatient studies with non-psychotic psychiatric patients. *Trans. 5th Ann. Res. Conf.*, VA, 1960, **5**, 101-106.
9. RICKELS, K., & DOWNING, R. Personality and attitudinal correlates of response to drug treatment in psychiatric outpatients: Theoretical orientation and measuring instruments to be employed. This journal.
10. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abnorm. & Soc. Psychol.*, 1953, **48**, 285-290.
11. WYLIE, R. C. The Self Concept. Lincoln: Univ. Nebraska Press, 1961, p. 370.

University of Pennsylvania  
School of Medicine  
Department of Psychiatry  
Philadelphia 4, Pennsylvania



## "HYPNOTIC BEHAVIOR" AS A FUNCTION OF TASK MOTIVATION\*

*State Hospital, Medfield, Massachusetts and Division of Psychiatry,  
Boston University School of Medicine*

THEODORE XENOPHON BARBER AND DAVID SMITH CALVERLEY

### A. INTRODUCTION

Weitzenhoffer and Sjöberg (25) and Barber and Glass (12) have recently presented similar experimental studies in which a series of standardized test-suggestions were administered to a group of Ss under a control and a "trance induction" condition. Comparable findings were obtained in the two experiments: (a) under the experimental condition ("trance induction procedure"), approximately half the Ss showed higher scores on the test-suggestions and the others showed no change in score or, in a few instances, a slight drop in score; and (b) as a group, the Ss showed a statistically significant enhancement of "suggestibility" following the "trance induction." These findings raise a basic question: Of the multiplicity of factors included in an experimental treatment termed a "trance induction," which are effective and which superfluous to facilitating response to suggestions?

The experimental treatment employed in the Weitzenhoffer and Sjöberg and Barber and Glass studies included the following factors, any one of which could possibly have produced enhanced "suggestibility":

1. The situation was explicitly defined as hypnosis and, presumably, implicitly defined by the connotations of the word "hypnosis" as different, as important, as a situation in which maximal response to suggestions was desired and expected.

2. Attempts were made to produce task motivation in the test situation; i.e., after "rapport" was established, the experimenter tried to allay the subject's apprehensions, to arouse his interest in the experiment, to enhance

\* Received in the Editorial Office on June 19, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by a grant (MY4825) from the National Institute of Mental Health, U.S. Public Health Service. The authors are indebted to Professor Kenneth Wolkon, President James Eteson, and Dean John Elberfeld of Worcester Junior College for making the facilities of the college available for the conduct of Experiments I-III; to Joan Burke for assistance in carrying out Experiment IV; and to the following for critically reading a preliminary draft of the manuscript: Maressa H. Orzack, Arje Latz, Paul D. Parker, and Harry Freeman.

his expectations that unusual and interesting events were forthcoming, and to facilitate his willingness to "try his best," "to let things happen," and "to cooperate."

3. The *S* stared at a fixed spot (Weitzenhoffer and Sjöberg) and listened to the monotonous sound of a metronome (Barber and Glass); remained relatively immobile for 15-20 minutes; while receiving repetitive suggestions that his eyes were closing and that he was becoming relaxed, passive, drowsy, and sleepy.

4. The *S* was explicitly or implicitly rewarded (by the intonations of the experimenter's voice or by other cues) for "appropriate" responses to the experimental instructions or suggestions.

5. The *S* was explicitly or tacitly rewarded for performing objectively "unusual" experimental behavior, for example, for behaving in a "drowsy" or "lethargic" manner.

Experiment I was designed to assess the relative effect of each of the above factors in producing enhanced "suggestibility." The hypotheses to be tested, derived from our previous studies (1-13, 17), are as follows:

*Hypothesis I.* Comparable levels of "suggestibility" are produced by a conventional trance induction procedure (Factors 1-5) and by brief instructions designed to produce task motivation in a test situation (Factor 2).

*Hypothesis II.* Comparable levels of "suggestibility" are produced by a conventional trance induction procedure (Factors 1-5) and by brief instructions designed to produce task motivation in a test situation *which is defined to the subject as "hypnosis"* (Factors 1 and 2).

*Hypothesis III.* Rewarding the *S* for correct responses on experimental tasks (Factor 4) facilitates response to subsequent test-suggestions.

*Hypothesis IV.* Rewarding the *S* for performing "unusual" behavior in an experimental setting (Factor 5) enhances response to subsequent test-suggestions.

## B. EXPERIMENT I

### 1. Method

*a. Subjects.* A total of 138 students (126 males and 12 females) from three sections of the Fall 1961 elementary psychology course at Worcester Junior College participated in the experiment in fulfillment of the course requirements. The students had not participated in previous experiments of this nature and had not had prior contacts with the experimenter.

*The Edwards Personal Preference Schedule* was first administered to



the three psychology sections. The students were then told that: the experiment involved imagination and perception as related to personality; and, to complete the next part of the experiment successfully, it was absolutely necessary that subjects not discuss it with each other.

The Ss were then randomly assigned by sex to one of six experimental groups, each of which contained 21 males and two females.

b. *Experimental conditions.* All Ss were exposed to the experiment individually and treatment sessions of the six types were randomized. The experimental treatments were as follows:

Condition 1, *Conventional Trance Induction Procedure.* Ss allocated to Group 1 were first told that they were to be hypnotized. They were then given the following instructions designed to produce task motivation in the hypnotic situation:

I am sure that you will find hypnosis a most interesting experience. Your ability to be hypnotized depends on your willingness to cooperate. It has nothing to do with your intelligence or your will-power. If you want to, you can pay no attention to me and remain awake all the time. In that case, you might make me seem silly but you are only wasting time. On the other hand, if you pay close attention to what I say, and follow what I tell you, you can easily fall into a hypnotic sleep and experience the interesting things I will tell you to experience. In this case you will be helping this experiment and not wasting your time. Hypnosis is nothing fearful or mysterious. It is merely a state of strong interest in some particular thing. In a sense you are hypnotized whenever you see a good movie and forget that you are part of the audience but instead feel a part of the story. Nothing will be done that will in any way cause you the least embarrassment. Your willingness to cooperate and your interest is what I ask for.

A standardized trance induction procedure was then administered as follows: the S was asked to fixate on a blinking light and to listen to the sound of a metronome; and received repeated suggestions of eye-closure, heaviness, relaxation, lethargy, drowsiness, and sleepiness, and suggestions that he would perform and experience everything that was to be suggested. During the induction procedure, the S was tacitly rewarded (by the intonations of the experimenter's voice) for complying with the suggestions.

Upon completion of the trance induction procedure, standardized at 15 minutes, eight test-suggestions which had been standardized and employed in previous studies (12, 17) were administered and scored as described below under *Criteria of Suggestibility*.

Condition 2, *Task Motivation in a Situation Defined as a Test of Imagi-*

nation. To structure the experiment as a test situation and to control the amount of time the *S* spent in the experimental setting, Group 2 *Ss* were first given two dittoed pages containing five Wechsler-Bellevue Verbal subtests and instructed to write the answers to the Wechsler items for a period of 15 minutes. During this period there was no conversation between the subject and the experimenter. After the Intelligence Test period was completed, the *S* was given the following instructions designed to produce task motivation in a situation defined as a test of imagination:

In the next part of the experiment I'm going to test your ability to imagine and to visualize. How well you do on the tests which I will give you depends entirely upon your willingness to try to imagine and to visualize the things I will ask you to imagine. Everyone passed these tests when they tried. For example, we asked people to close their eyes and to imagine that they were at a movie theater and were watching a show. Most people were able to do this very well; they were able to imagine very vividly that they were at a movie and they felt as if they were actually looking at the picture. However, a few people thought that this was an awkward or silly thing to do and did not try to imagine and failed the test. Yet when these people later realized that it wasn't hard to imagine, they were able to visualize the movie picture and they felt as if the imagined movie was as vivid and as real as an actual movie.

What I ask is your cooperation in helping this experiment by trying to imagine vividly what I describe to you. I want you to score as high as you can because we're trying to measure the maximum ability of people to imagine. If you don't try to the best of your ability, this experiment will be worthless and I'll tend to feel silly. On the other hand, if you try to imagine to the best of your ability, you can easily imagine and do the interesting things I tell you and you will be helping this experiment and not wasting any time.

Immediately following these instructions which subsumed a period of approximately one minute, the eight test-suggestions were administered and scored.

Condition 3, *Task Motivation in a Situation Defined as "Hypnosis."* Group 3 first wrote the answers to items from the five Wechsler-Bellevue Verbal subtests for a period of 15 minutes. Following the Intelligence Test period, this group was treated in the same way as Group 1 as described above, with the exception that the standardized trance induction procedure was *not* administered; that is, response to the test-suggestions was assessed immediately after instructions were given to produce task motivation in a situation defined to the subject as "hypnosis."

Condition 4, *Control*. The procedure for Group 4 was designed to assess



base-level response to the criterion suggestions. Ss in this group wrote the answers to the Wechsler subtests for a period of 15 minutes. Following this, they were told, "The next part of the experiment is a test of your imagination," and the test-suggestions were administered immediately.

Condition 5, *Reward Situation*. Ss in Group 5 answered the Wechsler items orally for a period of 15 minutes and were verbally rewarded by the experimenter ("that's very good") for correct responses. The S was then told that the next part of the experiment consisted of a test of imagination and the test-suggestions were administered immediately.

Condition 6, *Rewarding of "Unusual" or Role-Playing Behavior*. Subjects assigned to Group 6 were instructed as follows:

For the first part of this experiment I want you to try to imagine that you are a 6-year-old boy (or girl). You come from a middle-class family and you are in the first grade in school. (The experimenter further describes the child's home environment and relationships with parents and siblings.) Now try to the best of your ability to imagine that you are this child. Answer my questions as if you are this child.

During the following 5-minute period, the S was given items from the Wechsler subtests to answer orally and was verbally rewarded ("that's very good") for responses similar to those expected from a 6-year-old child.

The S was next instructed to imagine that he was a 20-year-old moron with a mental age of nine years; following this, he was instructed to imagine that he was a 96-year-old senile person. During this period (10 minutes) the S was first rewarded for responses to the Wechsler items which were similar to those expected from an adult with a mental age of nine and then rewarded for responses which resembled those expected from a senile person.

Upon completion of the 15-minute period of rewarded role-playing the S was told that the next part of the experiment also consisted of a test of imagination and the test-suggestions were administered immediately as described below.

c. *Test-suggestions (criteria of suggestibility)*. After each of the above experimental treatments was completed, the S was asked to close his eyes and was given eight test-suggestions:

1. Arm Lowering. Starting with the S's right arm extended and horizontal, suggestions are given for 30 seconds that the arm is becoming heavy and is moving down. Scoring criterion: one point for response of four inches or more.

2. Arm Levitation. Starting with the S's left arm extended and hori-

zontal, suggestions are given for 30 seconds that the arm is weightless and is moving up. Scoring criterion: one point for response of four inches or more.

3. Hand Lock. *S* is instructed to clasp his hands together tightly with fingers intertwined and place them in his lap. Suggestions are given for 45 seconds that the hands are like steel, they are welded together, they cannot be taken apart. Scoring criteria:  $\frac{1}{2}$  point for incomplete separation of hands after 5-seconds effort; one point for incomplete separation after 15-seconds effort.

4. Thirst "Hallucination." Suggestions of extreme thirst are given for 45 seconds. Scoring criteria:  $\frac{1}{2}$  point if *S* shows noticeable swallowing, moistening of lips, or marked mouth movements; additional  $\frac{1}{2}$  point if the subject indicates during the post-experimental interview that he became thirsty during this test.

5. Verbal Inhibition. Suggestions are given for 45 seconds that the *S*'s throat and jaw muscles are rigid and he cannot say his name. Scoring criteria:  $\frac{1}{2}$  point if the *S* does not say his name after 5-seconds effort; one point if he does not say his name after 15-seconds effort.

6. Body Immobility. Suggestions are given for 45 seconds that the *S*'s body is heavy and rigid and he cannot stand up. Scoring criteria:  $\frac{1}{2}$  point if subject is not completely standing after 5-seconds effort; one point if not completely standing after 15-seconds effort.

7. "Posthypnotic-Like" Response. *S* is told: "When this experiment is over I'll click like this (experimenter presents auditory stimulus) and you'll cough automatically. At the moment I click you'll cough." Scoring criterion: one point if the *S* coughs in response to auditory stimulus.

8. Selective Amnesia. *S* is told that when the experiment is over he will remember all the tests except the one where he was told that his arm was moving up (Test-Suggestion 2), and that he will remember this test only when the experimenter says, "Now you can remember." Scoring criterion: one point if the *S* does not refer to the amnesic task but recalls all other test-suggestions through Test-Suggestion 6, and then recalls Test-Suggestion 2 in response to the cue words.

Following the administration of the test-suggestions, the *S* was informed that the experiment was over and told to open his eyes. The auditory stimulus was then presented (to score Test-Suggestion 7); and the *S* was asked to state what had occurred (to score Test-Suggestion 8). The maximum score attainable on the test-suggestions was eight points.<sup>2</sup>

2. Copies of *The Scale of Eight Test-Suggestions* and standardization norms for the scale are available from the authors upon request.



*d. Subjective reports.* After the test-suggestions were scored, all *Ss* were asked the following questions and their replies recorded verbatim:

1. Which of these eight tests did you try to resist? (The experimenter describes each of the test-suggestions.)
2. Which of these tests did you actually feel? (The experimenter describes the test-suggestions that the *S* had passed with either  $\frac{1}{2}$  or 1 point.)
3. Which of these tests (test-suggestions passed) did you go along with just to follow instructions or to please me?
4. Which of these tests (test-suggestions passed) could you have resisted if you had really tried?

After the questions were answered, the *S* was admonished not to discuss the experiment with the other *Ss* and was then dismissed.

*e. Scoring criteria.* Question 1 was scored as follows: one point for each of the eight test-suggestions which the *S* said that he had attempted to resist.

The remaining three questions were scored with reference to those test-suggestions which the *S* had passed with a score of either  $\frac{1}{2}$  or one point, as follows:

Question 2: one point for each test-suggestion passed which the *S* claimed that he had "actually felt."

Question 3: one point for each test-suggestion passed which the *S* said that he went along with just to follow instructions or to please the experimenter.

Question 4: one point for each test-suggestion passed that the *S* said that he could have resisted if he had tried.

## 2. Results

Duncan's multiple range test (14) was employed to make multiple comparisons among the six means on: (a) scores on the eight test-suggestions combined; (b) scores on each of the test-suggestions taken individually; and (c) scores on each of the four questions (subjective reports).<sup>3</sup> The mean scores on each of these categories and the results of Duncan tests applied to the means are presented in Tables 1 and 2.

*a. Hypothesis 1.* The first hypothesis states that comparable levels of

---

3. A series of variance analyses was also performed on the scores. The results of these analyses paralleled the results of Duncan tests: (a) Scores on the eight test-suggestions combined differed significantly at the 1 per cent level of confidence ( $F = 7.84$  with 5 and 132 *df*). (b) With the exception of Test-Suggestion 8 (Selective Amnesia), scores on each test-suggestion differed at the 5 per cent level or better. (c) Scores on Questions 2 and 3 differed at the 5 per cent level but scores on Questions 1 and 4 were not significantly different.

"suggestibility" are produced by (a) a conventional trance induction procedure; and (b) brief instructions designed to produce task motivation in a test situation. This hypothesis was tested by comparing the scores of Group 1 (Conventional Trance Induction Procedure) with the scores of Group 2 (Task Motivation in a Situation Defined as a Test of Imagination). Table 1 indicates that Ss in both Group 1 and Group 2 passed, on the

TABLE 1  
MEAN SCORES ON TEST-SUGGESTIONS AND SUMMARY OF DUNCAN MULTIPLE  
RANGE TESTS APPLIED TO THE MEANS

Test-Suggestion	Group					
	1	2	3	4	5	6
1. Arm lowering	.739a	.652ab	.435bc	.304c	.261c	.261c
2. Arm levitation	.869a	.826ab	.609b	.391c	.348c	.348c
3. Hand lock	.935a	.804ab	.783ab	.804ab	.696b	.652b
4. Thirst "hallucination"	.848a	.913a	.804ab	.500b	.522b	.500b
5. Verbal inhibition	.869a	.739ab	.630abc	.674abc	.543bc	.478c
6. Body immobility	.891a	.674ab	.478b	.609b	.522b	.435b
7. "Posthypnotic"— like response	.696a	.609ab	.565abc	.478abc	.348bc	.217c
8. Selective amnesia	.391a	.391a	.261a	.217a	.130a	.174a
Total	6.24a	5.61ab	4.56bc	3.98cd	3.37cd	3.06d

NOTE.—Means containing a subscript letter in common are not significantly different from one another at the .05 level by Duncan multiple range test. Any two means which do not contain a common letter in the subscript differ significantly from each other at the .05 level by Duncan test.

average, approximately six of the eight test-suggestions. The results of Duncan tests applied to the means, as summarized in Table 1, indicate that the mean score of Group 1 on the total eight test-suggestions (6.24) does not differ significantly from the mean score of Group 2 (5.61); and the mean scores of the two groups do not differ significantly on any one of the eight test-suggestions taken separately. Table 2, which summarizes the results of Duncan tests applied to the four questions, indicates that the two groups differed in subjective reports, as follows: (a) the Trance Induction group stated more often than Group 2 (or any of the other groups) that it did not try to resist the suggestions and that it just went along with the suggestions to please the experimenter; and (b) the Task Motivation group stated more often than Group 1 (or any of the other groups) that it had "actually felt" the suggested effects.

The results thus fail to disconfirm Hypothesis I: it appears that a conventional trance induction procedure may be no more effective than brief



instructions designed to produce task motivation in a situation defined as a test of imagination in facilitating response to test-suggestions and in eliciting subjective reports that the suggested effects were experienced.

*b. Hypothesis II.* The second hypothesis states that comparable levels of "suggestibility" are produced by a conventional trance induction procedure and by brief instructions designed to produce task motivation in a test situation which is defined to the *S* as "hypnosis." This hypothesis was tested by comparing the scores of Group 1 (Conventional Trance

TABLE 2  
MEAN SCORES ON FOUR QUESTIONS (SUBJECTIVE REPORTS) AND  
SUMMARY OF DUNCAN RANGE TESTS APPLIED TO THE MEANS

Question	Group					
	1	2	3	4	5	6
1. Number of test-suggestions subject tried to resist.	1.69a	2.76b	2.43b	2.61b	3.04b	3.30b
2. Number of test-suggestions passed which were "actually felt."	3.83b	5.22a	3.35b	2.74bc	1.82c	1.78c
3. Number of test-suggestions passed which subject went along with just to follow instructions or to please the experimenter.	1.39a	.304b	.478b	.348b	.609b	.826ab
4. Number of test-suggestions passed which subject could have resisted if he had "really tried."	3.00a	2.87a	2.43a	2.26a	1.96a	2.13a

NOTE.—Means containing a subscript letter in common are not significantly different from each other at the 5 per cent level by Duncan multiple range test. Any two means which do not contain a common alphabetical symbol in the subscript differ significantly from one another at the 5 per cent level by Duncan test.

Induction Procedure) with the scores of Group 3 (Task Motivation in a Situation Defined as "Hypnosis"). The results of Duncan tests, as summarized in Table 1, indicate that: (a) the mean score of Group 1 on the eight test-suggestions combined (6.24) is significantly higher than the mean score of Group 3 (4.56); and (b) the mean scores of Group 1 are higher than the mean scores of Group 3 on each of the test-suggestions taken individually and are significantly higher on three of the eight suggestions. The results thus disconfirm Hypothesis II: it appears that brief instructions designed to produce task motivation in a situation defined as "hypnosis" are not as effective as a conventional trance induction procedure in eliciting enhanced "suggestibility."

A comparison of the scores for Groups 2 and 3 suggests the possibility that defining the situation as "hypnosis" may be less effective in facilitating response to suggestions than defining the situation as a "test of imagination." Groups 2 and 3 were given similar instructions to produce task-motivation in the experimental situation. However, to Group 2 the situation was defined as a "test of imagination" and to Group 3 the situation was defined as "hypnosis." Table 1 indicates that although the differences between the two groups fell short of statistical significance, the mean scores of Group 2 were higher than the mean scores of Group 3 on each of the eight test-suggestions.

Our unquantified observations of *Ss*' behavior under Conditions 1, 2, and 3 suggested the following possible interpretation of these findings: (a) when the situation is defined as "hypnosis" but a conventional trance induction procedure is not administered, resistance is aroused in some *Ss* who view "hypnosis" with anxiety or fear; (b) when the situation is defined as a "test of imagination" less resistance is aroused; and (c) when the situation is first defined as "hypnosis" and then a conventional trance induction procedure is administered, initial resistance aroused in some *Ss* by the connotations of the word "hypnosis" is abated by subsequent suggestions intended to produce relaxation, passivity, and drowsiness. Additional research is needed to quantify and to confirm these observations.

*c. Hypothesis III and IV.* Hypothesis III held that verbally rewarding the *S* for correct responses on an experimental task facilitates response to subsequent test-suggestions. Hypothesis IV stated that rewarding the *S* for carrying out "unusual" behavior in an experimental setting enhances performance on subsequent test-suggestions. These hypotheses were appraised by comparing the scores of Group 5 (Reward Situation) and Group 6 (Rewarding of "Unusual" or Role-Playing Behavior) with the scores of Group 4 (Control). The results of Duncan tests, as summarized in Table 1, fail to support these hypotheses: the mean scores on the total eight test-suggestions of Groups 4, 5, and 6 (3.98, 3.37, and 3.06, respectively) do not differ significantly from each other; and the mean scores of the three groups are not significantly different on any of the test-suggestions taken individually.

Table 1 also indicates that Control, Reward Situation, and Rewarding of Role-Playing Behavior were generally less effective in facilitating response to suggestions than Conventional Trance Induction Procedure, Task Motivation in a Situation Defined as a "Test of Imagination," and Task Motivation in a Situation Defined as "Hypnosis."



### 3. *Conclusions from Experiment I*

The above experimental findings suggest the following tentative conclusions:

1. Praising or verbally rewarding the *S* for correct responses on an experimental task or for "unusual" or role-playing behavior performed in the experimental setting does not seem to be a significant factor in enhancing performance on subsequent test-suggestions.

2. Instructions designed to produce task motivation in a situation defined either as a "test of imagination" or as "hypnosis" are effective in eliciting heightened suggestibility and heightened ability to perform behaviors which have been traditionally associated with the concept of "hypnotic trance."

3. A conventional trance induction procedure may be no more effective than brief instructions designed to produce task motivation in a situation defined as a "test of imagination" in facilitating response to test-suggestions and in eliciting subjective reports that the suggested effects were experienced.

### C. EXPERIMENT II

A major finding from the above experiment, that a conventional 15-minute trance induction procedure may be no more effective than 1-minute instructions designed to produce task motivation in a test-of-imagination situation in evoking enhanced response to test-suggestions, appears crucial to theory construction in the area of "hypnosis" and requires additional empirical test. The hypothesis was tested again in Experiment II.

#### 1. *Method*

- a. Subjects.* Thirty-six students (30 males and six females) enrolled in one section of the Spring 1962 introductory psychology class at Worcester Junior College served as *Ss* as part of their regular course requirements. The students had not participated in previous experiments of this nature and had not had prior contacts with the experimenter.

As in the preceding experiment, *The Edwards Personal Preference Schedule* was administered; and the class was told that: (a) the experiment was concerned with imagination, perception, and personality; and (b) to complete the experiment successfully it was necessary that *Ss* not discuss it with each other. The students were then randomly assigned by sex to one of two experimental treatments with 15 males and 3 females in each treatment.

- b. Experimental Conditions.* Two experimental conditions from the preceding experiment—Condition 1 (Conventional Trance Induction Procedure)

and Condition 2 (Task Motivation in a Situation Defined as a Test of Imagination)—were replicated with one change: the Wechsler subtests were not administered under Condition 2.

Under the Trance Induction condition Ss were first given instructions designed to produce task motivation in the hypnotic situation and were then given a standardized 15-minute conventional trance induction procedure. Under the Task Motivation condition subjects were not given a conventional trance induction procedure; instead, instructions designed to produce task motivation in a situation defined as a test of imagination were given for a 1-minute period. Response to the eight test-suggestions that had been used in Experiment I was again employed as the dependent variable. Subjective experiences were assessed by the same four questions used in the preceding experiment.

## 2. Results

Table 3 summarizes the results. As this table indicates: (a) S in both the Trance Induction group and the Task Motivation group passed, on the average, approximately five of the eight test-suggestions; (b) the two groups did not differ significantly in response to the eight test-suggestions combined and in response to any of the test-suggestions taken separately; (c) the two groups did not differ in response to any of the four questions (subjective reports).

The results are thus in line with the findings obtained in the preceding experiment. In both experiments, comparable responses to test-suggestions and comparable reports of subjective experiences were found after (a) a conventional trance induction procedure and after (b) brief instructions designed to produce task motivation in a situation defined to the subject as a test of imagination.

## D. EXPERIMENT III

In the above experiments Ss were given a series of eight test-suggestions and it appears possible that responses to the earlier suggestions in the series affected responses to the latter suggestions. Experiment III was performed to determine if instructions designed to produce task motivation in the test situation are as effective as a conventional trance induction procedure in enhancing response to a relatively "difficult" test-suggestion *given alone*. The test-suggestion selected for study consisted of suggestions of amnesia for experimental events.

### 1. Method

a. *Subjects.* Twenty-seven women enrolled in two sections of the Spring 1962 elementary psychology course at Worcester Junior College participated



in the experiment in fulfillment of the course requirements. These students had not participated in previous experiments of this nature and had not had prior contacts with the experimenter. They were told that the experiment involved imagination, perception, and personality. The *Ss* were randomly assigned to a Control condition (14 *Ss*) and a Trance Induction condition (13 *Ss*).

TABLE 3  
MEAN SCORES ON EIGHT TEST-SUGGESTIONS AND ON FOUR QUESTIONS (SUBJECTIVE REPORTS) WITH *t* TESTS FOR SIGNIFICANT DIFFERENCE BETWEEN GROUPS

Test-Suggestion	Group		<i>t</i> tests of difference
	Trance induction	Task motivation	
1. Arm lowering	.555	.722	1.030*
2. Arm levitation	.611	.444	.994*
3. Hand lock	.583	.750	1.114*
4. Thirst "hallucination"	.694	.861	.385*
5. Verbal inhibition	.611	.611	0*
6. Body immobility	.555	.583	.178*
7. "Posthypnotic"—like response	.500	.444	.328*
8. Selective amnesia	.444	.388	.332*
Total	4.55	4.80	.258*
<i>Question</i>			
1. Number of test-suggestions subject tried to resist.	3.39	3.55	.200*
2. Number of test-suggestions passed which were "actually felt."	3.78	4.11	.374*
3. Number of test-suggestions passed which subject went along with just to follow instructions or to please the experimenter.	.55	.44	.321*
4. Number of test-suggestions passed which subject could have resisted if he had "really tried."	2.39	2.94	.728*

\* Not significant

*b. Experimental conditions: Control condition.* To structure the experiment as a formal test situation, the Control group was given two dittoed pages containing five Wechsler-Bellevue Verbal subtests. *S* was instructed to write the answers to the Wechsler items for a period of 15 minutes. During this period there was no conversation between the *S* and the experimenter. Immediately following the Intelligence Test the *S* was given seven relatively meaningless tasks to perform as follows:

1. Hold your right arm in front of you. Now let it fall into your lap.

2. Raise your left arm as slowly as possible until it is pointed straight up. Now bring your hand to a normal position.
3. Squeeze your hands together tightly. Now release them.
4. Hold your mouth open and breathe through your mouth. Now close your mouth and breathe normally.
5. Say your name.
6. Stand up and then sit down.
7. Cough.

Upon completion of the tasks *S* was told:

In the next part of the experiment I'm going to test your ability to control your memory by determining how well you are able to forget something I will tell you to forget. Do you remember which shoe you put on first this morning? It is hard for everyone, including me, to remember how they put their shoes on. Most people can forget very easily what they did just a few days ago or even a few minutes ago. We've asked people if they remembered what they did one or two days ago at a certain time and a great majority couldn't remember off hand, unless they thought very hard and made an effort to remember. I often find it very hard to remember what I read in a newspaper or in a novel just a few minutes after I've read it. It really is very easy to forget some things.

As you know, I just gave you seven tasks to perform as follows: (1) to let your right arm fall; (2) to raise your left arm; (3) to squeeze your hands together tightly; (4) to breathe through your mouth; (5) to say your name; (6) to stand up; and (7) to cough.

In this next part of the experiment I want to determine how good you are at controlling your memory by forgetting the seven tasks which I just gave you. Now listen, this is very important. I want you to try to the best of your ability to erase from your memory all recollection of the seven tasks you have performed. Try to the best of your ability to forget the tasks. Forget them until I bring them back to your mind by saying, "Now you can remember." Now forget them completely.

The *S* was then given the Wechsler-Bellevue Information subtest; following this, she was asked, "What can you tell me about the experiment up to now?" After the *S* replied, she was asked, "Was there anything else?"; and then told, "Now you can remember everything perfectly well." *Ss* who had behaved as if they were amnesic for the tasks were asked two additional questions:

- A. Were you really unable to remember?
- B. How did you succeed in forgetting?
- c. *Trance Induction condition.* *Ss* assigned to the "trance induction" treat-



ment were first informed that they were to be hypnotized. Instructions were then given to alleviate apprehensions with respect to hypnosis and to produce personal involvement in the situation and positive motivation to perform well on the experimental tasks. A standardized 15-minute trance induction procedure was next administered which included eye-fixation on the blinking light of a metronome and repetitive suggestions of relaxation, feelings of heaviness, drowsiness, and sleepiness. Upon completion of the trance induction procedure, the *S* was instructed to perform the seven relatively meaningless tasks in the same way as the Control *Ss* as described above. The experimenter then spoke as follows:

I just gave you seven tasks to perform (the experimenter enumerates the tasks in the same way as with the Control group). Now listen, this is very important. You will erase from your memory all recollection of the seven tasks you have performed. You will forget all of the tasks. You will forget them until I bring them back to your mind by saying, "Now you can remember." You will forget them completely.

The *S* was awakened; the Wechsler Information subtest was administered; and replies were obtained to the same questions which had been asked the Control group.

## 2. Results

Three *Ss* (21 per cent) in the Control group and three subjects (23 per cent) in the Trance Induction group showed "symbolic amnesia," i.e., they did not recall any of the tasks when asked, "What can you tell me about the experiment up to now?", or when subsequently asked, "Was there anything else?" Each of these *Ss* then recalled at least six of the seven tasks when told, "Now you can remember everything perfectly well." When subsequently asked, "Were you really unable to remember?", these six subjects claimed that they had "actually forgotten."

The three "amnesic" *Ss* in the Control group gave the following answers to the question, "How did you succeed in forgetting?": "I don't know"; "I put them out of my mind by thinking of schoolwork and different things—that is, what I'll do tonight and tomorrow"; "I grouped the tasks together as 'a group' to be forgotten." The three "amnesic" *Ss* in the Trance Induction group answered this question as follows: "I don't know"; "I forgot them because you told me to forget them"; "I just kept saying 'forget' to myself."

The number of *Ss* in the Trance Induction group exhibiting apparent "amnesia" (three of 13 *Ss*) does not differ significantly from the number showing apparent "amnesia" in the Control group (three of 14 subjects) ( $\chi^2 = .13, p > .70$ ).

### 3. Discussion

The experimental results fail to reject the null hypothesis: it appears that suggestions of "amnesia" may be no more effective under a "trance induction" condition than under a control condition, provided that unselected Ss have been assigned at random to the two conditions and provided that the Control group is motivated to perform well on the experimental task.

It should be emphasized that these results were obtained with a homogeneous group of Ss who were randomly assigned to the hypnotic and control treatments. If, as in some previous hypnotic experiments, individuals rated as "good" hypnotic Ss had been assigned to the "trance induction" condition and individuals rated as "poor" hypnotic Ss or unrated Ss had been assigned to the control condition, a higher proportion of Ss in the "trance induction" group may have shown behavior indicative of "amnesia." However, such an experimental design confounds pre-existing differences in "suggestibility" with the treatment effects (8, 24). If "suggestible" individuals, that is, "good" hypnotic Ss, are allocated to the "trance induction" treatment, it is necessary, to control S differences, also to allocate "good" hypnotic Ss to the control treatment. If preexisting differences in "suggestibility" and motivational factors and interpersonal variables are controlled, the results of the present experiment strongly suggest that the null hypothesis of no difference in response to suggestions of "amnesia" in a "trance induction" and a control group will be difficult to reject.

It should also be noted that the "amnesia" exhibited in the present experiment consisted of what Hull (19) has termed "symbolic or verbal amnesia," that is: (a) Ss did not verbalize the "amnesic" tasks when asked by the experimenter to state what had occurred; (b) they verbalized the tasks when the experimenter said, "Now you can remember"; and (c) they later testified that they had "actually forgotten." It does not follow from this that "amnesia" would have been manifested if a person other than the experimenter had conducted the interviews, or if the experimenter had described the "amnesic" tasks and asked the Ss if they recognized them. However, it should also be noted that when investigators refer to "posthypnotic amnesia" as manifested by "deep trance" or "somnambulistic" hypnotic Ss they refer in practically all instances to a similar "symbolic or verbal amnesia" as was manifested in the present experiment. The evidence available at present, recently reviewed elsewhere (9), appears to indicate that "good" hypnotic Ss who show "complete verbal or symbolic amnesia" when interviewed by the hypnotist: (a) *recognize* the material which they claim not to remember;



and (b) characteristically show very little if any effects of the "amnesia" when tested by indirect methods which do not depend on verbal reports, such as assessment of practice effects or of retroactive inhibition effects. Furthermore, experimental studies have not been designed to determine if "good" hypnotic Ss who deny memory for experimental events when questioned by the hypnotist also deny memory for the events when interviewed by a prestigious person who is ostensibly not associated with the experiment but with whom Ss have a friendly or close relationship. Additional experiments are needed to test this possibility.

#### E. EXPERIMENT IV

Some months before the above experiments were performed, we had carried out an experiment to answer the following question: Are highly selected "deep trance" Ss more responsive than unselected control Ss to a "difficult posthypnotic-like suggestion," namely, to the suggestion to dream the following night on a selected topic? This earlier experiment differed from Experiments I-III, as described above, as follows: (a) selected "good" hypnotic Ss were assigned to the experimental group and unselected Ss were assigned to the control group; and (b) an earnest attempt was not made to motivate the control group to perform well on the suggested task. The experiment was performed as follows:

##### 1. Method

*a. Subjects.* Three groups participated: a group that was not given dream suggestions; a group given dream suggestions under control conditions; and a group given posthypnotic dream suggestions under "trance."

The Non-suggestion group and the Control group consisted of 35 and 11 women college students, respectively, who had not previously participated in experiments of this nature and had not had prior contacts with the experimenter. The Hypnotic group consisted of six women students; these Ss were selected from an original group of 72 women students at Regis College who had been tested for hypnotic susceptibility, as meeting criteria for "deep trance," showing such phenomena as hypnotically-induced negative and positive hallucinations, hypnotic age-regression, and posthypnotic amnesia.

*b. Experimental Conditions: Non-Suggestion Condition.* The 35 Ss assigned to this condition were told in a group setting that an experiment on dreams was being conducted. They were asked to watch carefully for any dreams that might occur during the following night and to record the dreams as soon as possible after their occurrence.

c. *Control condition.* The 11 control Ss were exposed to the experiment individually. To structure the experiment as a formal test situation, items from five Wechsler-Bellevue Verbal subtests were administered for a period of 15 minutes. The experimenter verbally rewarded ("that's very good") correct responses to the Wechsler items. Immediately after the Intelligence Test administration, S was told that the next part of the experiment consisted of dream study, and was given the following suggestion:

Tonight when you are in bed asleep you will have a dream about riding a bicycle. Sometime tonight when you are in bed asleep you will have a dream in which you will ride a bicycle. When you wake up in the morning write the dream on a piece of paper and bring it to me.

If the S asked any questions concerning the dream suggestion, she was told that no questions could be answered until the experiment was completed.

d. *Hypnotic condition.* The six hypnotic Ss were exposed to the experiment individually. To induce personal involvement in the situation the S was told that she was selected for the experiment because she possessed an ability which most people lacked—the ability to be deeply hypnotized—and that, possessing this ability, she would now have interesting experiences. A standardized 15-minute trance induction procedure was then administered which consisted of: (a) eye-fixation on a light blinking in synchrony with the sound of a metronome; (b) suggestions designed to produce task motivation in the hypnotic situation; and (c) repetitive suggestions of relaxation, drowsiness, and sleepiness. Suggestions were next given to induce body immobility, regression to age three, and negative and positive visual hallucinations. Each S appeared to enter "deep trance" and responded positively to the suggestions. Following this, the suggestion to dream the following night about "riding a bicycle" was given in the same way as with the Control group as described above. S was then awakened, nothing further was said concerning the posthypnotic suggestion, and S was dismissed.

e. *Subjective reports.* The following dittoed questionnaire was given the next day to those Ss presenting a dream concerned with "riding a bicycle":

This is a scientific study of dreaming which will be published in a scientific journal. Please answer the following questions as truthfully and objectively as you possibly can:

A. In what way did your dream about a bicycle differ from your other night dreams?

B. When did you dream it?; in the middle of the night?; in the morning before awakening?; etc.

C. Did you sleep normally? Did you wake up more than you usually do or sleep more fitfully? Did you have a harder time than usual in falling asleep?



## 2. Results

*a. Non-suggestion condition.* Of the 35 Ss who were asked to record their night dreams without receiving dream suggestions, seven submitted reports of dreams. None of these dream reports included a bicycle in the manifest contents.

*b. Control condition.* On the day following, nine of the 11 Ss who had received dream suggestions under non-hypnotic conditions reported either that they did not dream or that they could not remember their dreams. The other two Ss presented the following reports:

"I had a dream in which I was riding a bicycle. There were a lot of other people riding too and they were going up a hill."

"I went to bed at night and tried to dream about riding a bike. When I was almost asleep I started myself thinking about it and then different episodes that had occurred previously came to my mind—that is, a bike ride I once took with the girls."

In response to Item A on the questionnaire, both Ss wrote that their dream was similar to or differed only in content from other night dreams. In response to Items B and C, the first S wrote that she was not certain when during the night the dream appeared, and that she had slept normally except that she woke up during the night and thought about the dream suggestion; the second S reported that the "dream" appeared before going deeply asleep and she had slept normally.

*c. Hypnotic condition.* On the day following, four of the six hypnotic Ss reported that they did not dream or did not remember their dreams. The other two Ss presented the following reports:

"I dreamt I was getting on a bike but then the alarm rang and I woke up."

"The fact that I was to have a dream concerning me riding a bicycle must have made an impression on my mind. In a semi-conscious sleep I saw a tiny figure on a bicycle impishly fitting in and out as if it was going through an obstacle course."

In response to Items A and B on the questionnaire, both Ss said that the dream did not differ from other night dreams and it appeared in the morning before awakening. In response to Item C, the first S wrote that she "woke up all through the night and kept thinking I would have to get back to sleep and dream about riding a bike"; and the second S wrote that she had not slept normally—"I'm exhausted! I have the feeling of having been awake all night."

The number of Ss presenting "bicycle dreams" in the Hypnotic group (two of six Ss) does not differ significantly from the number reporting such "dreams" in the Control group (two of 11 Ss) ( $\chi^2 = 1.69$ ,  $p > .19$ ).

### 3. Discussion

Although only a small number of *Ss* were used in this experiment and the differences between the groups did not attain statistical significance, there appeared to be a trend toward greater responsiveness to the dream suggestion on the part of the Hypnotic group: 33 per cent (two of six) of the *Ss* in the Hypnotic group and 18 per cent (two of 11) in the Control group complied with the dream suggestion. It should be noted, however, that this experiment was deliberately designed to favor the Hypnotic condition as follows: (a) The Hypnotic group but not the Control group had participated in a previous selection experiment in which it had interacted with the experimenter and had an opportunity to form a friendly relationship with him. (b) The Control group consisted of unselected *Ss*. The Hypnotic group consisted of *Ss* meeting criteria for "deep trance"; this explicit criterion is difficult if not impossible to differentiate from an interrelated implicit criterion, namely, that *Ss* were highly responsive (with or without "hypnosis") to various types of suggestions. (c) Although the "trance induction procedure" administered to the Hypnotic group included instructions designed to produce task motivation in the test situation, the Control group did not receive motivating instructions. Additional experiments are required which control these factors. Hypnotic and control *Ss* should be randomly selected from a homogeneous group of "good" hypnotic *Ss* or from a homogeneous group of unrated *Ss* (8, 24). Both groups should have a comparable amount of prior contact and association with the experimenter. Both groups should be given motivating instructions prior to the administration of the dream suggestion. The findings from the present experiment appear to indicate that if *S* variables, interpersonal factors, and motivational variables are controlled, the null hypothesis, that hypnotic and nonhypnotic *Ss* do not differ in response to "posthypnotic-like" dream suggestions, will be difficult to reject.

The foregoing results also indicate the possibility that "suggested dreaming" may differ in essential respects from natural nocturnal dreaming. Hypnotic and control *Ss* presenting "dreams" on the selected topic appeared to show a conscious desire to execute the dream suggestion (e.g., "I went to bed at night and tried to dream about riding a bike"). Also, three of the four *Ss* who ostensibly complied with the dream suggestion stated either that their sleep was disturbed or that they awakened during the night and thought about the dream suggestion; and at least two strongly implied that they had purposely tried to "create" the dream (e.g., "When I was almost asleep I started myself thinking about it [the suggested dream topic] . . ."). In



a previous study Fisher (16) had similarly observed that some if not many Ss who complied with suggestions to dream on selected topics "had strong conscious desires to dream, would remind themselves of the suggestions during the day and before falling asleep, would awaken frequently during the night, think of the suggestion and make efforts to remember any dreams that they had."

Fisher's observations and the data obtained in the present experiment appear to be in line with other experimental findings, recently reviewed elsewhere (10), which suggest that: (a) "Dreams" presented by "deep trance" Ss in response to hypnotic or posthypnotic suggestions are produced when S is *awake* (according to electroencephalographic criteria), not during light sleep as appears to be the case during natural dreaming (23). (b) "Suggested dreams" may be functionally related to a factor which can be termed provisionally as "a striving on the part of the S to 'make' the dream appear." Additional experiments are needed to determine if "hypnotically-induced dreams" can be differentiated from "made-up dreams." A "deeply hypnotized" group should be given the suggestion to dream at night on a selected topic. A control group should be asked to make-up a dream-like narrative on the same topic. The findings from the present experiment and from Fisher's (16) experiment suggest the hypothesis that both groups will present similar "dream" reports.

#### F. GENERAL DISCUSSION

The results of the above experiments are in general agreement with the findings obtained in a series of other recent studies. Barber and Deeley (11) found that the performance which characterizes "deep trance" hypnotic Ss who have received suggestions intended to produce "color-blindness" can be elicited from control Ss who are instructed to try not to see the color red and the color green on the Ishihara plates. Oswald (21) and Barber (1) have presented data indicating that a performance that can be evoked from some "deeply hypnotized" Ss—"hallucinating" colors on a gray card, and subsequently reporting complementary colored afterimages (15, 22)—can also be evoked from some motivated control Ss by instructions to imagine and to visualize the specified colors.<sup>4</sup> Barber and Hahn (13) found that instructions to imagine vividly a pleasant situation when noxious stimulation was

4. Whether hypnotic Ss and control Ss who report complementary colored afterimages of "hallucinated" or imagined colors "perceive" the negative afterimages, or are able to verbalize the appropriate afterimages as a result of previous experiences with complementary colors, is an unanswered question that requires further research (cf., 18).

applied, given to appropriately motivated control Ss, were as effective as suggestions of analgesia given to "hypnotized" Ss in attenuating pain (produced by immersion of a limb in water at 2°C for three minutes) as indicated by subjective reports of reduced pain experience and by reduction in respiratory irregularities and forehead muscle tension. Also, a series of investigations, recently reviewed elsewhere (6), suggests the possibility that the behaviors and experiences subsumed under the terms "hypnotic blindness," "hypnotic deafness," and "hypnotically induced negative hallucinations" can be elicited from appropriately motivated control Ss by simple instructions to try to remain inattentive and unresponsive to visual or auditory stimulation. Along similar lines, London and Fuhrer (20) have presented data indicating that "hypnotized" Ss do not differ from motivated control Ss on measures of muscular strength and endurance. In brief, the results of Experiments I-IV and the results of other recent studies indicate that some if not many behaviors and experiences that have been traditionally termed "hypnotic" can be performed and experienced by Ss who have not received a conventional "trance induction procedure" and who do not appear to be "in trance," provided that the Ss are motivated to perform well on the criterion tasks.

#### G. SUGGESTIONS FOR FURTHER RESEARCH

The foregoing results suggest four major areas for further research:

1. Additional experiments are needed to compare a hypnotic group and a task-motivated group in performance of other behaviors which have been historically associated with the word "hypnosis" such as "hypnotic age-regression." With respect to "age-regression," the control group should be told in a convincing manner that previous experiences can be revived and relived by vividly imagining that one is at a former time. Following this, the control group should be given similar instructions as used in Experiments I and II to produce task motivation in the test-of-imagination situation. A series of studies reviewed elsewhere (5, 7) suggests that a Control group which has not been subjected to a procedure which has been traditionally termed a "trance induction" may not differ from a "Trance induction" group in response to suggestions of "age-regression" provided that the Control group has been motivated to perform optimally on the imaginative task and provided that disbeliefs concerning the possibility of experiencing this phenomenon have been minimized.

2. The task motivation treatment employed in Experiments I and II, which was found to have a strong effect in facilitating "suggestibility,"



included a number of independent variables. Under this condition Ss were told that: others had done well on the imaginative tasks; they also could perform well if they tried to imagine vividly; and if they did not try to the best of their ability to cooperate and to imagine, they would fail the tests, the experiment would be worthless, and the experimenter would tend to feel silly. It thus appears that the task motivation treatment included the following, and possibly other, factors: (a) competitive strivings to perform as well as others; (b) strivings to maintain or to enhance self-esteem; and (c) strivings to please or to fulfill the expectations of the experimenter. Additional experiments are needed to delineate the relative contribution of each of these variables to the treatment effect.

3. The standardized "trance induction procedure" employed in the above experiments included explicit instructions designed to produce task motivation in the hypnotic situation. It appears possible that this "induction" procedure (and other conventional "trance induction procedures") also included *implicit* task motivating factors; that is, the suggestions and manipulations which constituted the "trance induction" may have connoted the following to the Ss: "This is an unusual and important test situation in which you are expected to comply with all suggestions; if you do not try to the best of your ability to comply with the suggestions, you will disappoint the hypnotist, you will fail as an experimental S, and you will be hurting the experiment." The experimental findings presented above suggest that explicit and implicit task motivating instructions may be the effective factors in a "trance induction procedure" making for enhanced suggestibility. However, the findings do not exclude the possibility that other components of an "induction" may also have an effect. For instance, it appears possible that requiring the S to remain immobile with eyes closed for a period of time may itself facilitate response to suggestions. Additional experiments are required to determine if immobility, eye-closure, induced "drowsiness," and other factors subsumed under the term "trance induction procedure" contribute to the enhancement of suggestibility, or if these factors are irrelevant and explicit and implicit task motivating factors included in the "induction" are alone effective.

4. There is little reason to believe that the "trance induction procedure" employed in Experiments I-IV, which was standardized and administered in the same way to Ss allocated to the hypnotic treatments, was maximally effective. It appears possible that higher levels of performance may be elicited by varying the "trance induction" to "suit the S." However, there is also little reason to believe that the task motivation treatment administered to control groups in Experiments I and II had a maximal effect. Additional research

is needed to devise methods which are more effective than those used in the above experiments in producing optimal levels of task motivation among control *Ss*. The results of the foregoing experiments suggest that if optimal task motivating procedures are devised they may be found to be more effective than conventional "trance induction procedures" in eliciting "hypersuggestibility" and enhanced ability to carry out behaviors that have been historically associated with the word "hypnosis." If further research confirms this prediction, the following interesting eventualities may arise: (a) traditional "trance induction procedures" may no longer be employed to elicit "hypersuggestibility" and may gradually become historical curiosities; and (b) the concepts of "trance" and "hypnosis" may be replaced by the more general psychological concept of task motivation.

#### H. SUMMARY AND CONCLUSIONS

Four experiments are presented comparing the performance of a "trance induction" group with the performance of one or more control groups under varying levels of induced motivation.

In Experiment I an experimental group and five independent control groups were compared on response to the following eight test-suggestions: limb heaviness; limb levitation; inability-to-unclasp-hands; thirst "hallucination"; verbal inhibition; body immobility; selective amnesia; and "posthypnotic-like" response. The experimental group was assessed on the test-suggestions after receiving a conventional trance induction procedure. One control group was assessed on the test-suggestions without receiving preliminary instructions. The other four control groups were tested on the criterion suggestions under one of the following conditions: after receiving brief instructions designed to produce task motivation in a situation defined to the *S* as a "test of imagination"; after receiving instructions designed to produce task motivation in a situation defined to the *S* as "hypnosis"; after receiving verbal reward from the experimenter ("that's very good") for satisfactory responses on an experimental task; after receiving verbal reward for role-playing behavior performed in the experimental setting. The results indicated that: (a) Verbally rewarding or praising the *S* for "appropriate" responses on an experimental task or on role-playing tasks is not a significant factor in facilitating performance on subsequent test-suggestions. (b) Instructions designed to produce task motivation in a situation defined to the *S* either as a "test of imagination" or as "hypnosis" are effective in eliciting heightened suggestibility and heightened ability to perform "hypnotic-like" behaviors. (c) A conventional trance induction procedure may be



no more effective than brief instructions designed to produce task motivation in a test-of-imagination situation in producing enhanced ability to perform and to experience phenomena that have been historically associated with the word "hypnosis."

Experiment II replicated the major findings of Experiment I. Two groups of randomly selected Ss participated. One group was given instructions for a one-minute period designed to produce task motivation in a situation defined to the S as a test of imagination. The other group was given a conventional 15-minute trance induction procedure. Both groups were assessed on response to the eight test-suggestions that had been used in the preceding experiment. Comparable levels of overt response to the test-suggestions and comparable subjective reports that the suggested effects were experienced were found after the one-minute task motivating procedure and after the 15-minute conventional trance induction procedure.

In Experiment III unselected Ss were assigned at random to a control and a "trance induction" condition. The criterion task consisted of response to suggestions of amnesia for experimental events. An attempt was made to motivate the Control group to perform well on the criterion task. The two groups did not differ in response to the suggestion; a comparable proportion of Ss in both the Control group and the "Trance induction" group did not recall the experimental events until told, "Now you can remember," and subsequently testified that they had "actually forgotten."

Experiment IV assessed response to a "posthypnotic-like" suggestion to dream the following night on a specified topic in a selected group of "deep trance" Ss and in an unselected group of control Ss. In this experiment the "Trance" group received motivating instructions but the Control group did not receive such instructions. Although the differences between the groups did not attain statistical significance, there appeared to be a trend toward enhanced response to the dream suggestion on the part of the "deep trance" group.

The discussion relates these findings to other recent experimental studies, specifies the implications of the findings for further research, and draws the following conclusions:

1. Task motivation in the test situation appears to be a significant factor in producing "hypersuggestibility" or "ability to perform 'hypnotic-like' behaviors."

2. It appears possible that many if not all of the phenomena that have been historically associated with the word "hypnosis" can be elicited without

the administration of a procedure that has been traditionally termed a "trance induction" and without the Ss appearing to be "in trance," provided that Ss are motivated to perform well on the criterion tasks and provided that disbeliefs concerning the possibility of experiencing the phenomena have been minimized.

3. If these experimental findings are confirmed in replicative studies, the concepts of "trance" and "hypnosis" may no longer be useful in conceptualizing the factors involved in "hypersuggestibility" and performance of so-called "hypnotic behaviors"; response to "direct suggestions" or "commands" and performance of behaviors traditionally associated with the word "hypnosis" may be more parsimoniously conceptualized under the more general psychological concept of task motivation.

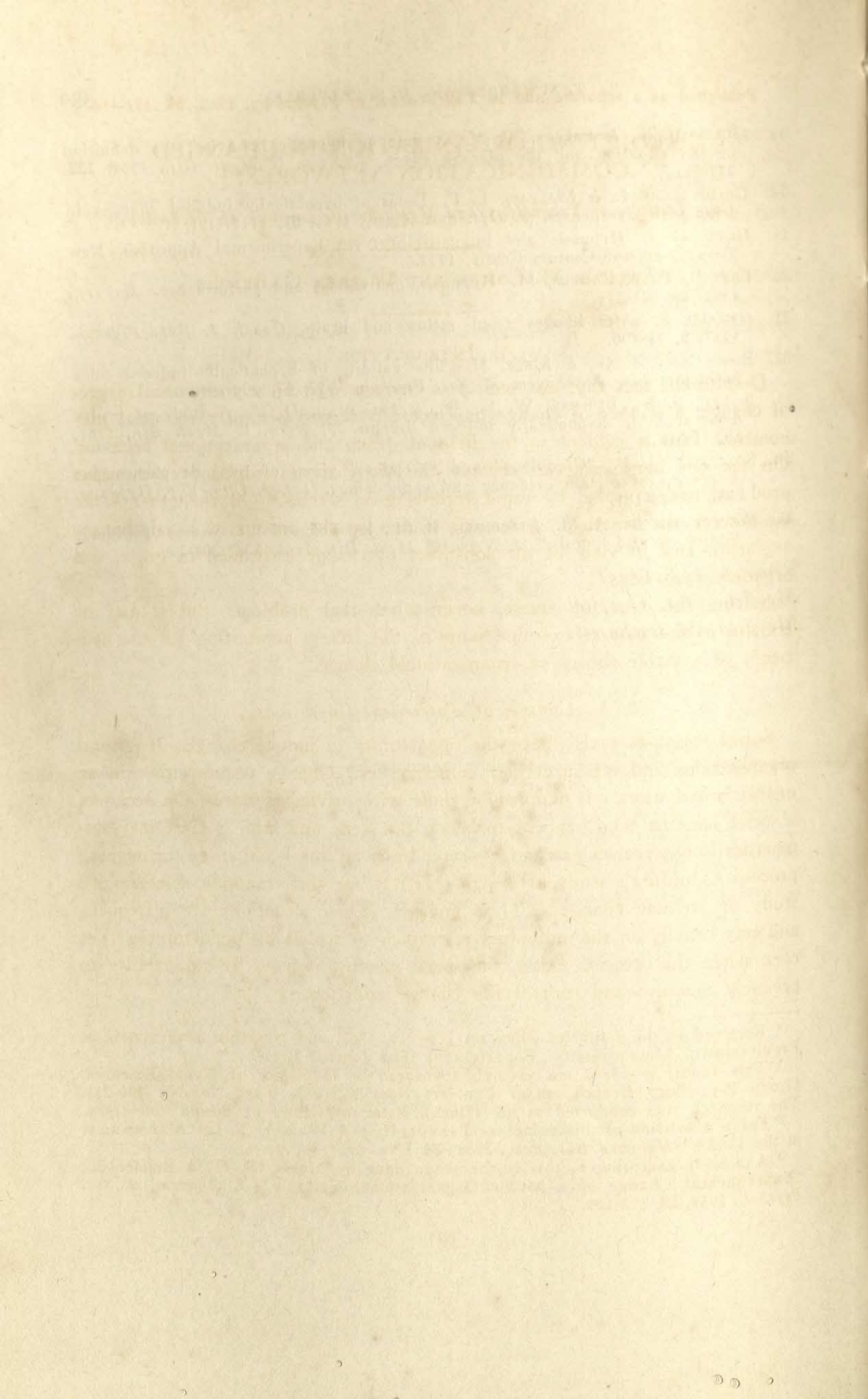
### REFERENCES

1. BARBER, T. X. The after-images of "imagined" and "hallucinated" colors. *J. Abnorm. & Soc. Psychol.*, 1959, **59**, 136-139.
2. ———. Toward a theory of pain: Relief of chronic pain by prefrontal leucotomy, opiates, placebos, and hypnosis. *Psychol. Bull.*, 1959, **56**, 430-460.
3. ———. "Hypnosis," analgesia, and the placebo effect. *J. A. M. A.*, 1960, **172**, 680-683.
4. ———. Antisocial and criminal acts induced by "hypnosis": A review of experimental and clinical findings. *Arch. Gen. Psychiat.*, 1961, **5**, 301-312.
5. ———. Experimental evidence for a theory of hypnotic behavior: II. Experimental controls in hypnotic age-regression. *Int. J. Clin. Exp. Hypnosis*, 1961, **9**, 181-193.
6. ———. Physiological effects of "hypnosis." *Psychol. Bull.*, 1961, **58**, 390-419.
7. ———. Hypnotic age-regression: A critical review. *Psychosom. Med.*, 1962, **24**, 286-299.
8. ———. Experimental controls and the phenomena of "hypnosis": A critique of hypnotic research methodology. *J. Nerv. & Ment. Dis.*, 1962, **134**, 493-505.
9. ———. Toward a theory of hypnosis: Posthypnotic behavior. *Arch. Gen. Psychiat.* (in press).
10. ———. Toward a theory of "hypnotic" behavior: The "hypnotically induced dream." *J. Nerv. & Ment. Dis.* (in press).
11. BARBER, T. X., & DEELEY, D. C. Experimental evidence for a theory of hypnotic behavior: I. "Hypnotic color-blindness" without "hypnosis." *Int. J. Clin. Exp. Hypnosis*, 1961, **9**, 79-86.
12. BARBER, T. X., & GLASS, L. B. Significant factors in hypnotic behavior. *J. Abnorm. & Soc. Psychol.*, 1962, **64**, 222-228.
13. BARBER, T. X., & HAHN, K. W., JR. Physiological and subjective response to pain under hypnotically-suggested and waking-imagined "analgesia" (to be published).
14. DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, **11**, 1-42.
15. ERICKSON, M. H., & ERICKSON, E. M. The hypnotic induction of hallucinatory color vision followed by pseudo-images. *J. Exp. Psychol.*, 1938, **22**, 581-588.
16. FISHER, C. Studies on the nature of suggestions: I. Experimental induction of dreams by direct suggestions. *J. Amer. Psychoanalyt. Assoc.*, 1953, **1**, 222-255.



17. GLASS, L. B., & BARBER, T. X. A note on hypnotic behavior, the definition of the situation and the placebo effect. *J. Nerv. & Ment. Dis.*, 1961, **132**, 539-541.
18. GOLDIAMDON, I., & MALPASS, L. F. Locus of hypnotically induced changes in color vision responses. *J. Opt. Soc. Amer.*, 1961, **51**, 1117-1121.
19. HULL, C. L. Hypnosis and Suggestibility: An Experimental Approach. New York: Appleton-Century-Crofts, 1933.
20. LONDON, P., & FUHRER, M. Hypnosis, motivation, and performance. *J. Pers.*, 1961, **29**, 321-333.
21. OSWALD, I. After-images from retina and brain. *Quart. J. Exp. Psychol.*, 1957, **9**, 88-100.
22. ROSENTHAL, B. G., & MELE, H. The validity of hypnotically induced color hallucinations. *J. Abnorm. & Soc. Psychol.*, 1952, **47**, 700-704.
23. SCHIFF, S. K., BUNNEY, W. E., JR., & FREEDMAN, D. X. A study of ocular movements in hypnotically induced dreams. *J. Nerv. & Ment. Dis.*, 1961, **133**, 59-68.
24. SUTCLIFFE, J. P. "Credulous" and "skeptical" views of hypnotic phenomena: A review of certain evidence and methodology. *Int. J. Clin. Exp. Hypnosis*, 1960, **8**, 73-101.
25. WEITZENHOFFER, A. M., & SJOBERG, B. M., JR. Suggestibility with and without "induction of hypnosis." *J. Nerv. & Ment. Dis.*, 1961, **132**, 204-220.

*Research Reprint*  
*Medfield State Hospital*  
*Medfield, Massachusetts*





## PREDICTING ORGANIZATION IN CHANGED COMMUNICATION NETWORKS\*<sup>1</sup>

*Case Institute of Technology and Massachusetts Institute of Technology*

---

ARTHUR M. COHEN AND WARREN G. BENNIS

---

### A. INTRODUCTION

Despite the fact that our society is characterized by a phenomenal degree of change, a science of change has woefully lagged behind the crucial phenomena. This is evident in the field of group and organizational behavior. On the one hand, we can observe the often acknowledged pre-occupation produced and revealed by rapid technological strides.<sup>2</sup> On the other hand, we observe the dearth of systematic studies on the process of social change; a glaring lack in view of the constancy of change embedded in group and organizational life.

Stating the issue, of course, never solves that problem. But it may be fruitful in this case to examine some of the factors accounting for the deficiency of a viable theory of organizational change.

#### 1. *Control of Antecedent Conditions*

Social scientists rarely have the opportunity to induce changes in natural organizations and measure their consequences.<sup>3</sup> Change occurs either in an unanticipated way or is planned by those with privileged status. On occasion, a social scientist who happens to be on the scene and who is alert may systematically observe such a change—serendipity willing—and study the ongoing process. Gouldner's study of a strike (7) is one such example. Lieberman's study of attitude change (12) is another. These situations are infrequent and rely heavily on the individual researcher to exploit his opportunities. Yet even when the occasion exists, the social scientist is beset by an inability to precisely measure and control the change-induction.

---

\* Received in the Editorial Office on June 19, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This report is based on research sponsored by the Office of Naval Research, Group Psychology Branch, under Contract Nonr 492(05), Task No. NR 170-255. The research was conducted at the Human Relations Center of Boston University.

<sup>2</sup> For a discussion of this point, see Leavitt, H. J. & Whisler, T. L., "Management in the 1980's," *Harvard Bus. Rev.*, 1958, 36 (No. 6).

<sup>3</sup> A notable exception to this is the study done by Morse, N. C., & Reimer, E., "Experimental Change of a Major Organizational Variable," *J. Abnorm. & Soc. Psychol.*, 1955, 52, 120-129.

## 2. *Lack of Systematic Data*

Do traces of the past co-exist with the present pattern or is the present a function of certain historical forces? This question echoes one of the chief doctrinal disputes in the social sciences between field theory and psychoanalytic theory, between competing models of economic growth and between different schools of social movements.<sup>4</sup> This issue is always capable of generating a heated exchange given the inadequacy of systematic data. The main difficulty has been the perceived inability to experiment with artificially induced historical data.<sup>5</sup>

## 3. *Time Unit*

Another deterrent to the systematic study of change processes has been the realistic problem of time perspective. The investigator has a limited period of time to study the phenomena, whereas change may unfold slowly, over a period of years, before the full effects could be shown and measured. This may account for the fact that most studies of change are historical in nature; that is, they deal with events which have taken place, which can be discussed from a vantage point of completion and separation.

The field of small groups, for example, which is characterized by a prolixity, a rigor, an experimental posture which has no equal in social psychology, with the possible exception of attitude measurement, has produced relatively very few studies on change. It seems impossible to simulate laboratory conditions which approximate change in natural settings. Most studies have been characterized by what one of the present writers (1) has called the "fifty-minute group." For the most part, this is not an exaggeration.

## 4. *Size and Generalizability of the Human Unit*

Most theories of social change deal with large collectivities, societies and cultures.<sup>6</sup> It is not obvious that the dynamics of change are similar for different orders of size. Until we have comparative data, the answer is not easily foreseeable. Furthermore those brilliant studies mentioned above reach levels of abstraction such that operational and empirical referents to other group studies cannot easily be made.

<sup>4</sup> For a discussion of this issue, see Moore, B., "Sociological Theory and Contemporary Politics," *Am. J. Sociol.*, 1955, **61**, 107-115.

<sup>5</sup> For a discussion of the application of the evolutionary model to the studies of linguistics, genetics and culture, see Gerard, R. W., Kluckhohn, C. & Rapoport A., "Biological and Cultural Evolution: Some Analogies and Exploration," *Behav. Sci.*, 1956, **1**, 6-34.

<sup>6</sup> D. Reisman's *The Lonely Crowd*, Yale Univ. Press, 1950; E. Fromm's *Man for Himself*, Rinehart, 1947; and A. Toynbee's *A Study of History*, Oxford Univ. Press, 1934-1959.



On the other hand, there exist numerous studies which do not suffer from excessive abstraction, but rather from restriction to their operational and empirical referents. These for the most part are excellent studies of change, providing rich clinical data about specific cases.<sup>7</sup> The study of change needs to include the positive qualities of both approaches referred to above. It needs to include descriptions of the properties of the situations under focus and classification of these properties so that greater abstractions can be made and generalizations validated.

All the foregoing has been said not to suggest that we can change defeat to victory, but to enumerate some of the obstacles preventing the development of more valid hypotheses concerning organizational change, which is the major interest of this paper.

Communication networks are suitable as a medium to begin such a task. The major substance of the relationships between members of any group is the communications among them. Everything must be done through some kinds of channels of communication, regardless of the contents of the tasks involved. A great many studies have been done using communication networks as experimental devices for contrasting the effects of different types of work structures on the behaviors of problem-solving groups.<sup>8</sup>

The initial impetus to these studies and probably the major justification as a procedure for studying groups in this manner was the belief that there were inducible topological properties of groups whose effects could be tested experimentally.<sup>9</sup> As often happens to a movement, once the primary creative impetus of the originators ceases, the original character of the movement is lost and new missions are adopted in a splintered fashion. In the case of studies of communication networks, the original Lewinian topological emphasis has disappeared. The early concern with establishing dimensions for comparing the structural properties of groups was disconnected from subsequent research interests. With a few exceptions, Shaw (14,

<sup>7</sup> Examples of this more clinical approach to change are: Jackson, J. M., "The Effect of Changing the Leadership of Small Work Groups," *Human Relat.*, 1953, 6, 25-44; Messinger, S. L., "Organizational Transformation: A Case Study of a Declining Social Movement," *Amer. Rev. Sociol.*, 1955, 20, 3-10; Richardson, S. A., "Technological Change: Some Effects on Three Canadian Fishing Villages," *Human Organ.*, 1952, 11, 17-27.

<sup>8</sup> For descriptions of most of the studies in this area, see Glazner, M., & Glazer, R., "Techniques for the Study of Team Structure and Behavior; Part II, Empirical Studies of the Effects of Structure," American Institute for Research, Pittsburgh, Pennsylvania, 1957.

<sup>9</sup> For a discussion of the application of Lewinian topological concepts to the study of group structures, see Bavelas, A., "A Mathematical Model of Group Structures," *App. Anthropol.*, 1948, 7, 16-30; also Bavelas, A., "Communication Patterns in Task-oriented Groups," *J. Acoust. Soc. of Amer.*, 1950, 22, 725-730.

15, 16, 17) in particular, research in the area of communication networks has not been systematic. One of the more promising series of efforts in this area has been connected with how groups organize for work in different communication networks (8, 9). Even these studies dealing with the organizing responses of people have only partly explored the potential of communication networks as a valuable testing ground for the study of groups. For example, studies of groups in communication networks have dealt with groups in static situations (one kind of network). This certainly in and of itself is not inappropriate. However, what has probably happened is that the case of studying groups in a single time slice has suppressed important latent interests in studying the unplanned activities of members of work groups under conditions of change. We certainly need descriptive generalizations about how people solve problems when they are in one kind of communication network as opposed to another. However, the focus of study, once such generalizations have been developed, needs to include concern for descriptions of how people react to changed situations.

This report consists of two experiments. Experiment I was a broad exploratory study in which a number of questions were investigated in addition to that of how groups organize for work when their communication networks are changed. Only the organizational findings will be reported here, since they provided the basis for deriving the organizational hypothesis tested in Experiment II. For a complete report of Experiment I, see (3).

#### A. EXPERIMENT I

##### 1. *Predicting Organization In Changed Communications Networks—I*

a. *Aims.* This study was intended to explore the effects of changes in the communication networks of problem-solving groups on efficiency, organization and satisfaction. Studies of communication networks have typically been of one hour's duration or less in which groups worked in one kind of communication network throughout the experiments. They have demonstrated that differences in efficiency, satisfaction and organization are produced by different communication networks.<sup>10</sup> These differences have been shown to be related to the kind of task used (10, 13, 16). For problems that are routine and repetitive, the more a communications network allows for equal participation, the higher is the satisfaction and the lower is the efficiency in times taken to solve problems and number of correctly solved problems. The effects of experience in more than one network have not been determined. The basis

<sup>10</sup> Op. cit. footnote 8.



for predicting the specific nature of the effects of different prior network experiences are equivocal and indirect. Because of this, three sets of predictions were used in this study, representing the logically possible outcomes of the effects of prior experience. The predictions compare the behaviors of groups that will have had experience in only one network with those of groups in an identical network but that will have had prior experience in a different network. The first set was based on the assumption that a difference in prior experience would not lead to differences in efficiency and satisfaction.<sup>11</sup> The second set was based on the assumption that the occurrence of a change in network, regardless of the kind of prior experience, would lead to higher efficiency and satisfaction.<sup>12</sup> The third set was based on the assumption that the particular kind of past experience would affect subsequent performance.<sup>13</sup> Prior experience in a more efficient network would lead to higher efficiency. Prior experience in a less efficient network would lead to lower efficiency. The same would be expected of satisfactions.

## B. METHOD

### 1. *Apparatus*

The apparatus used in this research was essentially the same as that used in the Leavitt study (11). The subjects were seated around a circular table so that each was separated by a vertical partition from the center to the edge of the table. At the center of the table was a five-layer pentagonal box with opening (slots) connecting all booths, and permitting subjects to push written messages to those with whom they were allowed to communicate. This apparatus allowed for the arrangement of any kind of communication network, simply by blocking up the appropriate slots. The networks referred to in this experiment and Experiment II are shown in Figure 1. Each booth had six mounted switches, each of which was labeled

<sup>11</sup> This assumption is consistent with that of the classical model of organization in which prior experience would be expected to play a minimal role in the outcomes of changes. When the properties of the new conditions are known and when the means of production and major relevant production contingencies are understood, there is assumed to be known all that is needed for control and prediction of the outcomes of changes.

<sup>12</sup> This assumption is based on the belief that breaking up the routine and monotony produced by continuously solving simple problems will improve performance and satisfaction. Alternate to this belief is the one in which change is believed to lead to resistance, and consequently, lower performance and satisfaction. The evidence for each belief is inconclusive. Preference dictated selection of the former.

<sup>13</sup> This assumption is consistent with that of the developmental model implicit in many psychological and sociological studies. It has the most face validity and common sense appeal.

and represented one of the possible answers. By pulling a switch, a subject automatically shut off a reaction timer and electrically registered his response on a master panel.

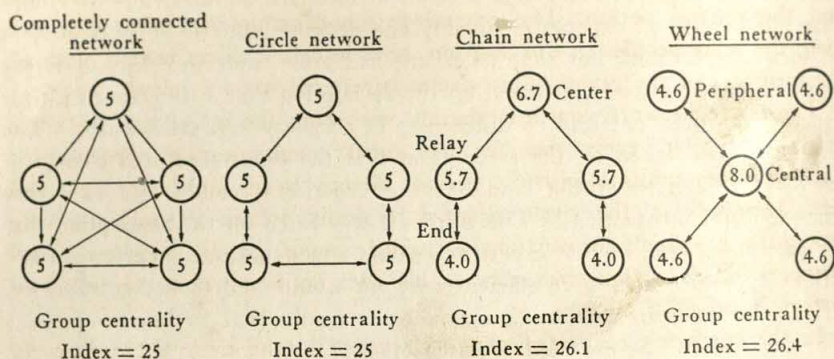


FIGURE 1  
 COMMUNICATION NETWORKS AND THEIR CHARACTERISTICS

The wheel network is considered the most structured and hierarchical on the basis of its centrality index (11). A group centrality index is the sum of the relative centrality indices of the positions. The centrality index of any position is defined as the ratio of the sum of the distances (number of steps to communicate) from all members to all members over the sum of the distances from a given position to all others. The higher the group centrality index the greater the inequalities in the communication opportunities of members in it. Thus, the wheel network represents the greatest degree of inequality. In this network, each of four members could communicate to the fifth (central position) person, but to no one else. The fifth person could communicate to everyone directly. The others had to go through him if they wanted to communicate with someone else. These four members were equal to each other in the communication restrictions imposed on them, but were much more restricted than the fifth member. The wheel typically follows a problem-solving procedure in which the four peripheral members send their information to the fifth or center member who decides on an answer and sends it back to them. This system is referred to as a central-hub.

The chain network is next highest in centrality, and consequently in restriction and communication inequality. In this network, there are two members who serve as end men, each of whom has one person with whom he can communicate directly. To this person the end men typically send information. Each of these persons serve as relay men, sending their informa-



tion along with that of their respective end men to the fifth person in the group who collates the information, decides on an answer and sends it to the two relay men, who in turn, send the answer to their respective end men. Thus, each relay man can communicate directly with two people: an end man and the central person. The central person communicates also with two men, the relay people. In this position, he is in closest direct contact with all the members of the group. This system is referred to as a relay.

The circle network contrasts sharply with both the wheel and the chain networks. In it, every member has equal communication opportunities. Each can communicate with the persons directly to his immediate right and left. Members in the circle network typically follow a problem-solving procedure in which information is passed around by all members to all members. Each person then acts as his own decision-making center. This system is called a circuit.

In the completely-connected network, just as in the circle network, every member has equal communication opportunities, hence the same group centrality index exists for both. But the completely-connected network, as opposed to the circle, imposes no communication restrictions on members. Members typically solve problems by passing and receiving information directly to and from others, then forming answers independently. This system is called an each-to-all.

## 2. Procedure

Two hundred subjects, paid volunteers, drawn from undergraduate classes at Boston University, were randomly assigned to forty groups of five men each. Ten groups were run in each of four experimental conditions. The four experimental conditions were: (a) Circle-wheel (c-w) in which groups first had to try to solve thirty problems in a circle network and thirty more in a wheel. This represented a change from a less to a more structured network. (b) Wheel-circle (w-c) in which groups had to try to solve thirty problems in a wheel network and then thirty more in a circle network. This represented a change from a more to a less structured network. (c) Circle-circle (c-c) in which groups had to try to solve sixty problems in a circle network. (d) Wheel-wheel (w-w) in which groups had to try to solve sixty problems in a wheel network.

Subjects, randomly assigned to positions (which were distinguished by having different colors) in their networks, were used only once. Each group spent from three to four hours in the experimental room. No subjects were used who had any knowledge of the experiment or who were color blind. Subjects were not informed of the kind of network in which they worked,

but could find out for themselves. By sending messages about who was communicating to whom, subjects were able to find out who was performing particular functions. Before actually solving the problems, the subjects read instructions which stressed the importance of working as a team as quickly as possible to get the correct answer. The procedure for each of the sixty problems was as follows: each subject was given a card, labeled by color and trial number, on which appeared a set of five out of six possible symbols. Each subject's card was different, in that the symbol lacking, the sixth one, was different in each case. Thus, for any given problem (called a trial) there was only one common symbol. The task was for each member to find the common symbol. Only written communication was allowed, between those who had open channels. There was no restriction on the amount of information in a message. Subjects were allowed to pass answers along. A trial was ended when all five subjects had registered answers. Subjects filled out questionnaires at the end of every fifth trial (five questions: certainty of answers, job satisfaction, existence of leadership, interpersonal satisfaction and interest in task). Each item had four alternatives, i.e., very certain of answers, somewhat certain of answers, somewhat uncertain of answers and very uncertain of answers.

Change was introduced after thirty trials, since previous research (2) had indicated that after thirty trials both wheel and circle conditions had stabilized the times taken to solve problems. The physical changes in communication channels were accomplished by substituting new pattern cards for the old ones which originally defined the communication networks (which blocked up various slots, and left others open in a systematic fashion). Changes were thus introduced without requiring any physical movements on the part of members.

### C. RESULTS

Only the results on the developments of problem-solving systems will be reported here, since these led to the hypothesis tested in Experiment II.<sup>14</sup>

#### 1. *Problem-solving Systems*

How groups organized for solving problems was determined by content analyses of message cards. A unit was defined as a sentence or a meaningful part of one. Kind and stability of organization were determined by coding information and answer exchanges. The procedure was similar to that used by

<sup>14</sup> Data about the information and answer exchange patterns in Experiment I are presented only for the w-c condition (Tables 1 and 2) which is an integral part of Experiment II. The kinds of problem-solving systems that were developed in all conditions (w-c, c-c, w-w, c-w, in Experiment I, and cc-cc and w-cc in Experiment II) are summarized in Table 8.



Guetzkow and Simon (9). The communication channels for each trial for each group were analyzed separately for information and answer flow. Two categories were used: the presence or absence of information and answer message units. The results of analyses in the separate trials were combined into blocks of five trials. The channels in each block were then sorted into three categories:

(1) *Column A* for channels that were *intermittently used*—at least once on one trial but less than all five trials of a block.

(2) *Column B* for channels that were *always used*—at least once on each of the five trials of a block.

(3) *Column C* for channels that were *never used*—not used even once on any of the five trials of a block.

For each group as the result of the above procedure there was obtained a series of trial blocks whose inspection enabled the experimenters to locate role differentiation where it did develop, on the basis of how positions in the communication networks were consistently and strategically placed in the information and answer exchange patterns. The results for each of the four experimental conditions is summarized as follows:

### 2. Information Exchange Patterns in c-c Groups

The information exchange patterns of the circles (c-c) started out as, and continued to be throughout, the kind in which each subject sent and received information from everyone with whom he was in contact.<sup>15</sup> The exchange patterns of information units revealed no role differentiation. There were neither centers for receiving nor special parts for sending in the circle groups. With very few exceptions, c-c groups throughout used all channels almost always for information exchanges.

### 3. Answer Exchange Patterns in c-c Groups

The answer exchange patterns of the circles were similar to their information exchanges. Communication channels were primarily intermittently used. Only infrequently were channels used always.

Problem-solving systems developed by the c-c groups can be summarized as follows: c-c groups conformed to a great extent to the operating limits of their networks. Each member used the channels available to him, continuously and in the same manner. Every member served as his own sending and re-

<sup>15</sup> A repeatedly supported observation of the networks we have studied is that the kinds of systems that ordinarily are most likely to develop in different kinds of networks are those that conform to the formal structures of the networks and tend to fill out the forms of the networks by utilizing all available channels.

ceiving station for processing information. Since practically all channels were used from the very beginning for this purpose, c-c can be considered to have stabilized very early in this respect. Answer exchanges, as might be expected, did not show a constant usage of channels, since members did not depend on others for answers but only for information and occasionally checking of answers. Thus, it was reasonable that channels used in answer exchanges would not have exhibited the constancy and stability of the information exchanges. Summing up, the c-c groups developed circuit problem-solving systems.

TABLE 1  
INFORMATION EXCHANGE PATTERNS FOR WHEEL-CIRCLE GROUPS (W-C) IN BLOCKS  
OF FIVE TRIALS FOR TRIALS 31-60

Group	Trial block																	
	7			8			9			10			11			12		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	7	2	1	2	4	4	0	4	6	0	4	6	0	4	6	0	4	6
2	4	5	1	2	5	3	4	4	2	1	4	5	0	4	6	0	4	6
3	2	8	0	2	4	4	1	4	5	0	4	6	0	4	6	0	4	6
4	4	6	0	0	5	5	1	4	5	0	4	6	0	4	6	0	4	6
5	7	3	0	3	3	4	1	4	5	0	4	6	0	4	6	0	4	6
6	1	8	1	0	8	2	2	8	0	0	8	2	0	8	2	1	7	2
7	4	4	2	0	4	6	0	4	6	0	4	6	0	4	6	0	4	6
8	3	7	0	3	7	0	0	7	3	3	6	1	1	6	3	0	6	4
9	3	7	0	2	6	2	3	4	3	0	4	6	0	4	6	0	4	6
10	0	10	0	2	8	0	3	5	2	1	4	5	0	4	6	0	4	6

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.

#### 4. Information Exchange Patterns in w-c Groups (see Table 1)

The information and answer exchange patterns of the circles that came from wheels (w-c) were extremely different from those of the groups that had always been circles (c-c). None of the c-c groups and eight out of the 10 w-c groups developed into perfect relay systems,  $P < .01$  using Fisher's exact test. A relay system typically develops in a chain network and is one in which one of the communication links of the circle network is unused, and in which three different kinds of roles develop: (a) a center man who receives information from the person to his right and the person to his left (both called *relay men*); (b) two relay men each of whom receives information from one man (not the center); and (c) two end-men each of whom sends information to and receives answers from one relay man. Thus, in the w-c groups there developed special parts for sending and receiving information. Data supporting such statements come from the following: of the 10 w-c



groups, eight developed perfect relay systems by the beginning of trial block 11 (Trial 21). No channels were used intermittently (zeroes in trial block 11, Column A). Four channels were always used (4's in Column B, trial block 11) (two end-men sent to two relay men who sent to the center man). Six channels were never used (6's in the Column C, trial block 11). Although all eight perfect relay systems did not develop before trial block eleven, they had developed either relay or relay-like problem-solving systems by the beginning of trial block 10 (Trial 46). Six were perfectly developed relay systems and two were relays except that the communication link between end-men was used intermittently and extraneously.

TABLE 2  
ANSWER EXCHANGE PATTERNS FOR WHEEL-CIRCLE GROUPS (W-C) IN BLOCKS  
OF FIVE TRIALS FOR TRIALS 31-60

Group	Trial block																	
	7			8			9			10			11			12		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	9	1	0	2	4	4	2	4	4	1	4	5	0	4	6	0	4	6
2	7	1	2	3	4	3	1	4	5	0	4	6	0	4	6	0	4	6
3	5	2	3	2	4	4	0	4	6	0	4	6	0	4	6	0	4	6
4	10	0	0	2	4	4	0	4	6	0	4	6	0	4	6	0	4	6
5	7	1	2	4	2	4	3	3	4	2	4	4	0	4	6	0	4	6
6	3	3	4	8	0	2	8	0	2	2	4	4	2	4	4	2	4	4
7	0	4	6	1	3	6	0	4	6	0	4	6	0	4	6	0	4	6
8	9	0	1	5	2	3	1	4	5	6	2	2	1	4	5	1	4	5
9	3	3	4	5	2	3	1	4	5	0	4	6	0	4	6	0	4	6
10	8	2	0	4	4	2	3	3	4	1	4	5	1	4	5	1	4	5

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.

##### 5. Answer Exchange Patterns of the w-c Groups (see Table 2)

The answer exchange patterns of the w-c groups paralleled those of the information exchange patterns. Seven of the ten w-c groups developed into perfect relay systems. Two of the remaining three w-c groups developed into almost perfect relays. The communication link between end-men was intermittently and extraneously used (it did not interfere with relay operation). The answer exchange patterns were fully developed by the beginning of trial block 11. As in the case of the information exchange patterns, the answer exchange patterns of the relay system had for all apparent purposes been developed by the beginning of trial block 10, Trial 46.

To sum up for the information and answer exchange developments of the w-c groups: Groups that were placed in circle networks, but which had had

problem-solving experience in wheel networks, developed relay problem-solving systems characteristic of chain networks. These relay systems were characterized by having three kinds of differentiated roles: (a) center men who collected information from and sent the answers to relay men, (b) relay men who had collected information from and sent answers to end-men; and (c) end-men who just sent information.

#### 6. *Information Exchange Patterns in w-w Groups*

Groups exchanged information in the following way: the four peripheral members sent their information to the central member. He did not pass his own information along to the others. This central-hub type of information exchange pattern had developed by the third trial block.

#### 7. *Answer Exchange Patterns in w-w Groups*

Groups developed the following kind of answer exchange pattern: the central member passed on the answers to the four peripheral members after having collected information from them. This central-hub pattern was developed by the third trial block.

#### 8. *Information and Answer Exchange Patterns in c-w Groups*

Groups in the c-w condition developed the same kind of information and answer exchange patterns as those which had always been in wheel networks, namely central-hub systems.

### 9. *Summary of Results*

The results of analyses on the developments of problem-solving systems (and other results—see (3) for a complete report) support the assumption that the particular kind of prior network experience would affect subsequent performance. Prior network experience influenced how members organized for solving problems, how quickly and correctly they solved problems, how certain they were of their answers, how satisfied they were and whether or not leadership was continuous from one network to another.

### D. DISCUSSION

The organizing responses of groups in new networks were influenced by their organizational experiences in prior networks. Training was transferred when the new network permitted it. Circle networks permitted such transfer. Wheels did not because of their high restriction. Nearly all of the groups in the wheel-to-circle condition developed relay systems, which were the



most centralized ones that could have emerged within the circle network. Wheel groups in the circle-to-wheel change showed no comparable development of systems atypical for the wheel network. The wheel network restrictions greatly reduced the likelihood of such systems developing. In the case of the w-c condition, the development of the more efficient relay systems led to shorter times taken to solve problems and more correctly solved problems than c-c groups. Since groups in the c-w and w-w conditions developed identical and typical problem-solving systems, their times taken to solve problems and numbers of correctly solved problems did not differ.

In addition to utilizing knowledge derived from experience<sup>8</sup> in prior networks, groups behaved rationally. When they were confronted with changed networks, from which more than one problem-solving system could have been developed, they selected procedures from those with which they had experience and which were logically the best for solving problems. When groups were changed from wheel to circle networks they initially gravitated toward the development of the circuit system which is typical for the circle network. This was only incompletely developed and was subsequently rejected. They became progressively more efficient utilizing the relay system typical of the chain network. Groups seemed to have responded to problem-solving cues that revealed the inadequacy of the circuit system—longer times and greater uncertainty of answers. Such cues stemmed from the comparatively greater efficiency of the central-hub systems used in the prior wheel networks—shorter times and greater certainty of answers. In order for w-c groups to have developed relay systems they had to impose severe operating restrictions on themselves. They spent significantly more of their effort on understanding network structure and on organizational planning than did c-c groups which did not develop role-differentiated problem-solving systems. In the w-c groups that developed relay systems there were members who organized activities more than others and who became leaders.

### *1. Changes in Communication Network and Emergent Problem-Solving Systems*

Within the limits of our type of experimental situation, we conclude that when a work structure is changed from a more to a less restricted one, people develop, on their own, efficient problem-solving; provided, the high restrictions characterizing such systems appear to be best for solving the problems at hand and provided the restrictions are self-imposed. When people (see (6) for a detailed discussion of this statement) find themselves in a condition in which their accustomed procedures for coping with required

environmental demands are no longer usable, they try to reduce the uncertainty of the new work situation by setting up new work procedures. The new procedures become adopted through efforts on the part of group members to apply what they had been doing in their past work structures; to apply a procedure of centralizing the work flow, for example, as was the case in Experiment I. Such procedures are not drafted in a wholesale and uncritical fashion. For one thing, the parameters of the new work structure might not permit this to be done. Thus, an approximation would have to be made. In our research, as an example, approximation to a system that is as centralized as possible, within the limits of the new structure (within the circle network), leads to the development of a relay system which is the most centralized system possible within the circle network. Secondly, rational contrast of previously used procedures and those possible within the new network (discovered primarily by trial and error organizational efforts), might reveal that such past procedures are no longer useful or are less useful than the newer ones made possible by the new work condition. If this were the case then they would not tend to be adopted. This was not the case in our experiment, however. Wheels which were antecedent to circle networks were empirically more efficient for solving problems of the kind used in our research.

#### A. EXPERIMENT II (see (6) for a separate report)

##### 1. *Predicting Organization in Changed Communication Networks—II*

In Experiment I, when wheel communication networks were changed to circles, relay problem-solving systems emerged. This finding was replicated as part of a study reported elsewhere (5). An explanation was offered above for the emergence of such systems. Although it has been established that this organizational event does occur when the networks used are wheels and circles, additional experimentation is needed to evaluate the explanation of it. If the explanation is valid, then we should be able to generate predictions about the kinds of problem-solving systems that would emerge when different kinds of networks are involved in the changes. Experiment II represents an attempt to do this by testing an hypothesis about emergent problem-solving systems in completely-connected networks (see Figure 1) which were changed from wheels.

Studying the effects of changes from wheels to completely-connected networks satisfies the necessary requirements for partially testing<sup>16</sup> the appro-

<sup>16</sup> Even with this additional experiment, the explanation can only be considered to have been partially tested. The changes in both Experiments I and II are samples



priateness of the explanation given in Experiment I. The wheel network is the most restricted of all possible two-way channel, five-man networks. The completely-connected network is the least restricted. The problem-solving system that develops in the wheel is more efficient<sup>17</sup> than the one typically developing in the completely-connected network. The change from wheel to completely-connected network allows for the most extreme test of predictions about the class of change conditions represented by the change from wheel to circle (see footnote 16).

It was concluded in Experiment I that change from a more restricted (centralized) and more efficient communication network leads to the development of as centralized a problem-solving system as possible within the limits of the new network. The completely-connected network allows for the development of all systems, the most centralized of which is the kind that develops in the wheel network (called a central-hub system) in which the four peripheral members send their information to the fifth and central member who collates the information and sends them answers. The problem-solving system (called an each-to-all system) that develops in the completely-connected network is one in which each member sends information to all others and forms his own answers.

## 2. Hypothesis

When wheel networks are changed to completely-connected networks, groups will develop problem-solving systems like those in the wheels—*central-hub systems*. The prediction that a central-hub system will be developed is based on the expectation that the principles of co-ordination utilized in the wheel network will be fully transferred, since the completely-connected network would allow for complete transfer and since the central-hub system would represent a more efficient system than the each-to-all one typically developing in the completely-connected network, for the particular class of problems.

---

of one class of change conditions, in which the prior network is more restricted (more centralized) and more efficient for solving the kind of problems presented to the groups. There are other classes of change conditions. These will be discussed at the end of this paper.

<sup>17</sup> This assumption was based on extrapolation from Leavitt's original Experiment (II), where networks could be ranked according to their efficiency on the basis of the amount of communication restrictions on members (see Figure 1). As a check, a "t" test was done in Experiment II on the times taken to solve problems by groups in wheel and completely-connected networks. Groups in wheel networks were more efficient, taking less time to solve problems.

## B. METHOD

### 1. *Apparatus*

The apparatus used was the same as in Experiment I.

### 2. *Procedure*

One hundred male subjects, paid volunteers drawn from undergraduate classes at Boston University were randomly assigned to 20 groups of five men each. Ten groups were used in each of two conditions: (a) groups tried to solve 60 problems in a completely-connected network; and (b) groups tried to solve 30 problems in a wheel network and then 30 more problems in a completely-connected network.

The remaining procedures were the same as in Experiment I.

## C. RESULTS

The procedures for analyzing the developments of problem-solving systems were the same as in Experiment I.

### 1. *Information Exchange Patterns in cc-cc Groups* (see Table 3)

The developments in the exchange patterns in the cc-cc condition include the whole range of trials (see Tables 3 and 4). This is so because the comparisons between groups in the cc-cc and w-cc conditions must include initial developments, which in the cases of the cc-cc groups start at Trial one, and in the case of w-cc groups start at Trial thirty-one.

Of the 10 groups, eight developed complete or near-complete each-to-all systems, (Groups 2, 3, 5-10). A complete each-to-all system would appear in the tables for any given trial block, as: 0 20 0 (for Columns A, B and C respectively). As Table 3 reveals (trial block 12), only one of these eight groups, Group 10, had systematically eliminated any channels from usage (had channels that were never used (Column C)). The great majority of the channels were always used (Column B), with the remainder of the channels being used intermittently (Column A). The total score of communication channels for any group for any trial block for the cc-cc network is twenty. Groups one and four developed central-hub systems for exchanging information. Pictorial representations of the communications histories of these two groups (not shown in the tables) revealed that leadership in these groups was unstable in that it shifted from one member to another; in the case of Group one, from Blue to Red; in the case of Group four, from Yellow to Blue. In each of these shifts, a number of trials (usually a full trial block) was taken to make the transition.



TABLE 3  
INFORMATION EXCHANGE PATTERNS FOR COMPLETELY-CONNECTED GROUPS (CC-CC) (In Blocks of Five Trials for Trials 1-60)

Group	Trial block											
	1			2			3			4		
	A	B	C	A	B	C	A	B	C	A	B	C
1	3 17 0	4 16 0	5 15 0	15 4 1	11 4 5	1 4 15	8 0 12	0 4 16	0 4 16	0 4 16	1 4 16	0 4 16
2	8 12 0	5 15 0	1 19 0	0 20 0	0 20 0	0 20 0	0 20 0	0 20 0	0 20 0	0 20 0	0 20 0	0 20 0
3	0 20 0	1 19 0	1 19 0	0 20 0	0 20 0	0 20 0	0 20 0	1 19 0	0 20 0	0 20 0	2 18 0	0 20 0
4	12 8 0	11 6 3	5 4 11	2 4 14	0 4 16	0 4 16	10 0 10	0 4 16	0 4 16	0 4 16	0 4 16	0 4 16
5	10 10 0	5 15 0	7 13 0	0 6 14	5 14 1	2 18 0	4 15 1	7 13 0	5 15 0	6 14 0	10 10 0	5 15 0
6	9 9 2	9 9 2	8 10 2	11 8 1	10 10 0	9 11 0	13 7 0	8 11 1	6 14 0	8 12 0	12 8 0	10 10 0
7	7 13 0	4 16 0	6 14 0	5 15 0	3 17 0	6 14 0	6 14 0	6 14 0	4 16 0	4 16 0	3 17 0	3 17 0
8	4 16 0	0 20 0	2 18 0	1 19 0	2 18 0	1 19 0	1 19 0	2 18 0	2 18 0	2 18 0	2 18 0	1 19 0
9	6 14 0	6 14 0	7 13 0	5 15 0	5 15 0	5 15 0	4 16 0	3 17 0	0 20 0	0 20 0	0 20 0	0 20 0
10	8 12 0	6 14 0	8 12 0	7 12 1	8 12 0	3 16 1	7 13 0	9 11 0	7 13 0	8 11 1	6 12 2	2 15 3

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.

TABLE 4  
ANSWER EXCHANGE PATTERNS FOR COMPLETELY-CONNECTED GROUPS (CC-CC) (In blocks of Five Trials for Trials 1-60)

Group	Trial block														
	1			2			3			4			5		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	19	0	1	17	0	3	16	0	4	10	0	10	2	4	14
2	18	0	2	20	0	0	15	0	5	4	0	16	11	0	9
3	16	1	3	12	0	8	9	0	11	8	0	12	6	0	14
4	16	3	1	8	3	9	7	3	10	1	5	14	1	5	14
5	19	0	1	19	1	0	16	4	0	17	3	0	14	0	6
6	8	1	11	10	2	8	13	2	5	15	4	1	9	5	6
7	13	1	6	8	3	9	18	0	2	14	3	3	15	2	3
8	15	5	0	19	0	1	20	0	0	16	4	0	14	2	4
9	16	1	3	19	1	0	17	0	3	16	1	3	19	0	1
10	18	2	0	19	0	1	16	1	3	19	0	1	19	1	0
Trial block															
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															
26															
27															
28															
29															
30															
31															
32															
33															
34															
35															
36															
37															
38															
39															
40															
41															
42															
43															
44															
45															
46															
47															
48															
49															
50															
51															
52															
53															
54															
55															
56															
57															
58															
59															
60															

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.



## 2. Answer Exchange Patterns for cc-cc Groups (see Table 4).

For those eight groups which developed each-to-all information exchange systems, the great majority of the channels were either never used (Column C) or were intermittently used (Column A) for exchanging answers. Since each individual served as his own data processing machine, exchanges of answers served primarily to hasten tardy members or for confirmation of results, but were not essential to the functioning of the system (group members did not need this to obtain their answers). It may be noted that Group two, by the last trial block, dispensed with answer exchanges altogether. Groups one and four which developed central-hub information exchange systems, were required to use a similar centralized system for answer exchanges, since only one person, the leader, derived the answers. This results in a pattern in which no channels are used intermittently (Column A), four channels are always used (Column B) and sixteen channels are never used (Column C trial block 12). An examination of pictorial representations (not shown in the tables) revealed that while Group one developed a central-hub answer exchange system, Group four developed an answer exchange system that typically accompanies a chain network, i.e., a system in which two members assist the leader in transmitting answers.

TABLE 5  
INFORMATION EXCHANGE PATTERNS FOR WHEEL-COMpletely-CONNECTED GROUPS  
(W-CC) (In Blocks of Five Trials for Trials 31-60)

Group	Trial block											
	7			8			9			10		
	A	B	C	A	B	C	A	B	C	A	B	C
1	1	4	15	0	4	16	0	4	16	1	4	15
2	8	4	8	0	4	16	0	4	16	1	4	15
3	15	4	1	9	11	0	14	5	1	12	8	0
4	6	4	10	0	4	16	0	4	16	0	4	16
5	5	4	11	0	4	16	0	4	16	0	4	16
6	5	4	11	0	4	16	0	4	16	0	4	16
7	10	4	6	6	6	8	2	4	14	2	4	14
8	6	4	10	0	4	16	0	4	16	0	4	16
9	12	3	5	0	4	16	0	4	16	0	4	16
10	5	4	11	0	4	16	0	4	16	0	4	16

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.

## 3. Information Exchange Patterns for w-cc Groups (see Table 5)

Nine of the 10 groups which had had prior experience in the wheel network, with its central-hub system, developed central-hub systems within

their completely-connected networks. Once the hub of the central-hub system was established, the person serving this function continued to do so throughout the trials. This contrasts with the instability of leadership in the two cc-cc groups that developed central-hub information systems. The pattern (fully developed by the beginning of trial block 11) of zero channels used intermittently (Column A), four channels used always (Column B) and 16 channels never used (Column C), reflects the use of the central-hub system. One group (number three) did not organize into any clearly defined system. Many channels were always or intermittently used and few channels were never used.

TABLE 6  
ANSWER EXCHANGE PATTERNS FOR WHEEL-COMpletely-CONNECTED GROUPS  
(W-CC) (In Blocks of Five Trials for Trials 31-60)

Group	Trial block											
	7			8			9			10		
	A	B	C	A	B	C	A	B	C	A	B	C
1	2	4	14	0	4	16	0	4	16	0	4	16
2	0	4	16	0	4	16	0	4	16	2	4	14
3	20	0	0	16	4	0	14	6	0	7	13	0
4	0	4	16	0	4	16	0	4	16	0	4	16
5	0	4	16	0	4	16	0	4	16	0	4	16
6	0	4	16	0	4	16	0	4	16	0	4	16
7	10	0	10	2	4	14	0	5	15	0	4	16
8	3	4	13	0	4	16	0	4	16	0	4	16
9	8	1	11	1	4	15	1	4	15	0	4	16
10	2	3	15	1	4	15	1	4	15	0	4	16

A = Channels used *intermittently*.

B = Channels *always* used.

C = Channels *never* used.

#### 4. Answer Exchange Patterns in w-cc Groups (see Table 6)

The nine groups which organized into central-hub information exchange systems utilized similarly centralized systems for the distribution of answers. This is reflected in the 0 4 16 pattern of channel usage (corresponding to Columns A, B and C respectively, trial block 12). Pictorial representations (not shown in the tables) revealed that Groups two and ten developed slightly modified central-hub answer exchange systems in which a relay assists the leader in the distribution of answers (the leader formed the answers in all nine w-cc groups referred to above). Group three, which did not organize a clearly defined system of information exchange, developed no clearly defined answer exchange system. This is reflected in the large number of channels intermittently used.



Fisher's Exact Test revealed that w-cc and cc-cc networks differed significantly in the kinds of information and answer exchange systems that groups in them adopted for solving problems (see Table 7). Cc-cc groups, without prior experience in wheel networks, did not develop central-hub problem-solving systems, but tended to utilize each-to-all systems. The reverse was true of the organizational behaviors of groups working in completely-connected networks, but which had had experience in wheel (w-cc) networks (and consequently had had experience with the central-hub problem-solving systems that groups in wheel networks develop). These w-cc groups developed central-hub problem-solving systems.

The problem-solving systems that developed in both Experiments I and II clearly reveal the influences of prior network experiences. Groups with experience only in circle networks did not develop the more efficient centralized system of the relay. Groups with experience only in completely-connected networks did not develop the more efficient and centralized central-hub system.

Two conditions existed which allowed past network experience to have these effects: (a) the antecedent networks were more efficient, and (b) the subsequent networks (especially so in the completely-connected network) allowed modification of communication channels to take place within their limits, so that transfer could take place and members could establish restrictions themselves. In the case of the c-w change (Experiment I) these two conditions were not met. Thus, no atypical problem-solving systems emerged in wheels that had been changed from circles.

#### D. DISCUSSION

The results, supporting the prediction that groups which have had prior experience in a wheel network will organize into central-hub problem solving systems when placed in a completely-connected network, enables us to discuss more justifiably and in more general terms the circumstances surrounding the emergence of different kinds of problem-solving systems.<sup>18</sup>

We have observed that groups tend to utilize all the communication channels available to them when there has been no alternative problem-solving experience. The more a problem-solving system diverges from that expected

<sup>18</sup> We have treated the issue of predicting the development of problem-solving systems primarily as a cognitive one. In so doing, we have assumed sufficient motivation on the part of each member. We have also tacitly assumed that ordinary individual idiosyncratic tendencies would not have profound effects on the development of problem-solving systems. However, from the fact that not all groups in the same conditions developed identical systems, we may infer that knowledge of idiosyncratic factors might help to account for exceptions to regularities in organizational developments.

TABLE 7  
COMPARISONS OF ORGANIZATIONAL BEHAVIORS BETWEEN COMPLETELY-CONNECTED (C-C) AND WHEEL-COMpletely-CONNECTED (W-CC) NETWORKS

	Information exchange		Answer exchange		
	Each-to-all system	Central-hub system	Each-to-all system	Central-hub system	Other (chain, modified, central-hub, etc.)
Completely-connected (CC) trials 1-60	8	2	8	1	(1)
Wheels-completely-connected (W-CC) trials 31-60	1	9	1	7	(2)
Fisher's exact test	$p = .01$		$p = .01$		



on the basis of full usage of a network's channels, the more we might infer the existence of deliberate choice on the part of group members. This is especially true when the problem-solving system that develops is the most extreme opposite to that which would be expected to develop from full usage of the available communication channels.

Our experimental situation illustrates this point. The development of a central-hub problem-solving system, instead of an each-to-all system within the completely-connected network, is the most extreme divergence (maintaining two-way connections) that can occur. Such a divergence requires that each of four members restrict his participation from contact with everyone and from being an equal, independent decision-maker, to just providing information to a fifth person serving as leader, who in turn provides answers. This voluntary transformation from a non-centralized system (each-to-all) to the most centralized (central-hub) system is the product of deliberate choice from alternatives that are perceived as unequal in efficiency. Past experience provides the opportunity for the existence of such alternatives. Logic provides the method of choice. The need to reduce ambiguity (efficiently find answers to problems) together with the opportunity to self-establish restrictions provides the motivation.

Past experience (transfer to training), logic (rationality) and the need to reduce ambiguity obviously require a concrete situation in which to operate. In this experiment, and others (3, 5), the prior experiences of groups were in networks whose problem-solving systems were more efficient than those of the new networks to which they were changed, i.e., in this study they were wheel networks in which central-hub systems develop. This means that we cannot separately evaluate the contributions of transfer of training and rationality to the development of problem-solving systems, unless we design a situation in which the past experience is with a system (in a network) which is less efficient than the system ordinarily developing in the new network. In such a case, we might expect that as the difference in efficiency increases in the direction which we have specified above, rationality, as the more dominant force influencing the development of problem-solving systems, would take precedence over transfer of training. The more the discrepancy becomes salient, the more likely will the past system be rejected in favor of the one ordinarily developing within the new network. This description represents a second class of change situations which needs testing. In this, the system that typically develops in the prior network is less efficient than that which typically develops in the subsequent network. The experimental conditions would be the same as in the test of the first class, but the task would

be changed so that the more centralized a system is the lower would be its expected efficiency. Under these conditions, if we were to change a wheel network with its central-hub system to a completely-connected network with its expected each-to-all system, we would expect the each-to-all system to be used.

There is another class of change situations which also needs testing. In it, the system that typically develops in the antecedent network is more efficient than that typically developing in the subsequent network, but less efficient than the most efficient system possible in the subsequent network. Testing this class: when we change from a chain network with its expected relay system to a completely-connected network with its expected each-to-all system, we would expect a central-hub system to be adopted. This prediction is based on an assumption in which, under conditions described above, a problem-solving group is expected to do more than transfer a previously adopted system. It will develop a problem-solving system in the most advanced form possible, utilizing the approach to problem-solving underlying the previously adopted system. In this case (where a chain is changed to a completely-connected network), this would be illustrated by the development of the most centralized system possible, the central-hub.

The construction of a viable model of change, whether it focuses on the effects of change on the development of problem-solving systems in communication networks, with which we have been concerned, or on other effects in other kinds of situations, requires that we: (a) identify characteristic ways people respond to classes of problem-solving situations, which themselves are; (b) classifiable according to distinguishable properties; so that (c) reliable predictions can be made when characteristic human response tendencies are applied to the characteristics of the specific situations.

In a limited way we have begun to do this. We have, of necessity, restricted our efforts thus far. We have chosen as our subject a very limited area of study, communication networks in small artificial groups, in which change is externally imposed without advance preparation of members. In addition, we have limited our attention to one-change conditions. What the cumulative effects of more than one change might be cannot be securely predicted from our research alone. Such predictions probably require the use of additional assertions about the nature of serial changes.

#### E. SUMMARY OF EXPERIMENT II

This study was intended to test the hypothesis that when wheel networks are changed to completely-connected networks, groups will develop problem-



solving systems that are identical to those ordinarily developing in wheel networks, i.e., central-hub systems, in which each of four members sends his information to a fifth who collates it and sends answers. This hypothesis was developed from an experiment (Experiment I) which was part of a series of studies (2, 3, 4, 5, 6) on the communication systems established by problem-solving groups.

In a Bavelas-Leavitt arrangement, ten groups of five men each were assigned to each of two experimental conditions: (a) a completely-connected condition in which members had to try to solve a total of sixty problems; and (b) a wheel-completely-connected condition, in which members had to try to solve thirty problems in a wheel and then thirty more problems in a completely-connected network. For all groups for all problems, the task was the same: to try to find the symbol common to all five members from six possible symbols. Problem-solving systems were measured by analyzing the content and direction of flow of messages sent by members.

The results revealed that groups in wheel-completely-connected networks developed central-hub problem-solving systems. Groups working only in the completely-connected network developed each-to-all and not central-hub systems.

It was concluded that when groups work in communication networks which have been changed from one kind to a distinct other, the kinds of problem-solving systems that they develop are influenced in the following way: the selections from the possible systems are different, through the systematic use of alternative experiences, derived from the past. Such systematic use is made of alternative experience derived from the past when: (a) the problem-solving systems adopted in the antecedent networks are more efficient than the ones most typically developing in the subsequent networks; and (b) when the subsequent networks allow for the specialized use of channels in them so that procedures of problem-solving can be transferred from the past and can be established by the members themselves.

#### REFERENCES

1. BENNIS, W. G. A Critique of Group Therapy Research. *International Journal of Group Psychotherapy*, Jan., 1960.
2. COHEN, A., BENNIS, W., & WOLKON, G. The effects of continued practice on the behaviors of problem-solving groups. *Sociometry*, 24, No. 4, Dec. 1961.
3. ———. The effects of changes in communication networks on the behaviors of problem-solving groups. *Sociometry*, Vol. 25, No. 2, June, 1962.
4. ———. The effects of position changes in communication patterns on organization and satisfaction. Boston University Human Relations Center Technical Report, 1960.

5. COHEN, A., & BENNIS, W. Continuity of leadership in communication networks. *Hum. Relat.*, **14**, No. 4, November, 1961.
6. COHEN, A., BENNIS, W., & BRIGGS, D. A Partial Test of a Model Predicting the Influence of Previous Experience on the Communication Systems Established by Problem-solving Groups. Boston University Human Relations Center Technical Report, 1960.
7. GOULDNER, A. Wildcat Strike. Ohio: Antioch Press, 1954.
8. GUETZKOW, H., & DILL, W. Factors in the organizational development of task-oriented groups. *Sociometry*, 1957, **20**, 233-250.
9. GUETZKOW, H., & SIMON, H. The impact of certain communication nets in task-oriented groups. *Mgmt. Sci.*, 1955, **1**, 233-250.
10. HEISE, G. A., & MILLER, G. A. Problem-solving by small groups using various communication nets. *J. Abnorm. & Soc. Psychol.*, 1951, **46**, 327-335.
11. LEAVITT, H. J. Some effects of certain communication patterns on group performance. *J. Abnorm. & Soc. Psychol.*, 1951, **46**, 38-50.
12. LIEBERMAN, S. The effects of changes in roles on the attitudes of role occupants. *Human Rel.*, 1956, **9**, 385-402.
13. MACY, J. JR., CHRISTIE, L. S., & LUCE, R. D. Coding noise in a task-oriented group. *J. Abnorm. & Soc. Psychol.*, 1953, **48**, 401-409.
14. SHAW, M. Group structure and the behavior of individuals in small groups. *J. of Psychol.*, 1954, **38**, 139-149.
15. ———. Some effects of unequal distribution of information upon group performance in various communication nets. *J. Abnorm. & Soc. Psychol.*, 1954, **49**, 547-553.
16. ———. Some effects of problem complexity upon problem solution efficiency in different communication nets. *J. Exper. Psychol.*, 1954, **48**, 211-217.
17. SHAW, M., & ROTHCHILD, G. H. Some effects of prolonged experience in communication nets. *J. Appl. Psychol.*, 1956, **40**, 281-286.

Case Institute of Technology  
Department of Management  
University Circle  
Cleveland 6, Ohio



## MEASUREMENT OF THE PRIMARY FACTORS OF THE TOTAL PERSONALITY\*<sup>1</sup>

*New York University*

---

THOMAS N. JENKINS

---

The initial incentive for this investigation was a study carried out in 1941-2. The writer assembled items from all the personality tests in existence both here and abroad. When this chore was completed, we had<sup>o</sup> over 7000 items, each typewritten on a separate card. The object of this investigation was to determine which trait categories were covered by these items. Content analysis was to be utilized first, followed by a study of the statistical relationships between item sets based on content analysis.

When we classified these items on the basis of content, we found that we were dealing with a relatively small number of categories. It was obvious that these test constructors failed to develop measures covering the entire personality sphere. It seems that the test makers throughout the world, influenced by contemporary theories in psychology, psychoanalysis and psychiatry, were all attempting to measure the same things, such as, extraversion, introversion, neuroticism, inferiority complexes and so on. Hence, we decided to initiate a project, the purpose of which was to determine the primary personality traits of the total personality.

### A. THE PLAN FOR A COMPREHENSIVE SURVEY OF PERSONALITY TRAITS

Attention should first be directed to two basic features of this project. (a) The investigation was largely exploratory. Modern psychologists tend to utilize hypothetico-deductive methodology for research. This procedure enables one to verify theories or hypotheses but tends to limit the research worker in a scientific journey of discovery. Logically, exploration precedes verification of hypotheses.

Exploration, if helter-skelter or non-methodical, is likely to be abortive. What is needed is methodical exploration followed by analytical verification by means of valid quantitative methods.

---

\* Received in the Editorial Office on June 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The publication of this study was facilitated by a grant from the Andreas Foundation, Minneapolis, Minn.

(b) Another important feature of the present investigation was complete reliance on the responses of subjects instead of the use of external criteria, such as personality ratings by others. At this point, let us examine the global personality study carried out by Cattell (7, 8). In this important pioneer investigation, Cattell carefully selects a list of 171 trait terms, which he organized into bipolar clusters. This check list was used to rate about 200 male subjects, divided into 13 groups. Each group was rated by two of their fellows. These ratings were factor analyzed.

It is worthy of note that Cattell in his global study made use of the opinions (ratings) of others instead of *actual* responses on the part of the subjects. One of the chief objections to this approach is that personality ratings of subjects by laymen are likely to be strongly colored by cultural bias. This point will be elaborated more fully later in the discussion of the pitfalls of content analysis. A similar error is likely to arise when a psychologist constructs a test based on expert opinion as an external criterion. If psychiatrists, for instance, rate subjects, their opinions are apt to be heavily colored by current theories regarding abnormal behavior.

For these reasons, the *responses* of subjects were analyzed instead of people's *opinions* about the subjects. In the absence of external criteria, such as ratings, the responses of subjects were correlated and analyzed. The factorial criterion was always used for final validation of the scoring categories for primary personality traits.

### 1. *Initial Phases of our Exploratory Approach*

We decided to gather items from as many different human sources as feasible. It was found that items of the sentence type could be supplied most readily. These sources comprised graduate students, undergraduate students, clinical psychologists, teachers, as well as laymen with no psychological background. Later study of these items showed that items furnished by people with very little psychological background proved to be highly effective in the discovery of primary personality qualities. This fact may be the result in part of the lack of theoretical psychological bias of lay contributors. These contributors merely indicated how they felt, what they imagined, or what they did under ordinary living conditions. A lay contributor, for instance, might state simply the circumstances under which he feels or dreams that he is floating. From a set of 2756 items of the declarative sentence type, 1204 items were selected for an experimental inventory. The response scale for this inventory will be described in the next section.



## 2. Attempts to Ascertain a Satisfactory Form of Response Scale

A careful study was made of student reactions to various forms of response scales. The simplest scale, of course, is a *yes-no* or *true-false* scale. The chief difficulty with this over-simplified response scale is that it fails to provide the test taker with the *relative* truth or falsity of the statement as it applies to him. The same would be true of a frequency scale of *always-never*. Test makers sometimes introduce a neutral category in the scale to satisfy the examinees. The scale, for instance, might be *true-can't say-false*.

We therefore made one study of student reactions to a neutral category in the response scale. It was found that examinees took advantage of the neutral category for various reasons, as for example, evasiveness and dubiousness. For instance, an individual who felt very confident would return fewer neutral responses than one who tended to feel very doubtful about everything. In this sense, the neutral category merely served as a very crude measure of confidence. For these reasons the neutral category was discarded in the development of our response scales.

The alternative to a neutral category was a broad response scale. We found that students reacted most favorably to such scales because the examinee could pinpoint his response for accuracy more exactly. There is also less feeling of response restriction. A study was also made of student reactions to frequency scales and quantitative response scales. This investigation clearly suggested that it was easier for the examinee to make a judgment in terms of frequency (how often?) than in terms of trait loading (how much?). The frequency response scale is especially satisfactory for the lower levels of intelligence.

We finally decided to utilize a six-point, closed-frequency scale, ranging from an upper limit of *always* to a lower limit of *never*. The closed type of scale tends to minimize the effect of one type of variance, namely, subjective notions concerning the width of the scale. The scale consisted of letters which stand for the following: *always, almost always, often, sometimes, hardly ever, and never*. The form of the item and scale is shown in the following example: A AA O S HE N 317. Eating between meals makes me feel calm.

## B. SOURCES AND TREATMENT OF EXPERIMENTAL TEST DATA

Our preliminary work was always conducted with student populations. Later, additional verification was sought through the use of diverse samples, such as industrial employees, retail saleswomen, military cadets and basic airmen. During the war it was difficult to obtain adequate samples of

students because the young men had entered military service. It seemed to the writer that students immediately after the war were more cooperative than during normal times in participating in programs of this sort. The magnitude of the program was one which was likely to exhaust the patience of both faculty and students.

The experimental inventory of 1204 items was given to approximately 800 undergraduate students who were taking courses in psychology, economics and history at New York University. An experimental test of the adjectival type also was administered. It comprised 685 adjectives concerning behavior and 59 referring strictly to the physical characteristics of the human body. The results of this adjectival test will not be discussed in this article. Our experience to date, however, indicates that the sentence type is more satisfactory than the adjectival form of item.

Two short published tests were also given to this student sample. One series was Forms *J* and *K* of the C-R Opinionnaire of Lentz (16). The other was the "generalized morale scale" of Rundquist and Sletto (18). These tests served as external criteria for general morale and conservatism.

Several procedures were found to be very useful and often essential for the selection and validation of items for the measurement of primary personality traits. Among these were the following:

1. Content analysis as a means of obtaining preliminary sets of items to be subjected later to rigorous methods of analysis.
2. Cluster analysis of item responses. A method for clustering was evolved which produced uniformly consistent results. Later this procedure was expanded into a system of contour analysis. The items were grouped on the basis of levels of correlation, so that several sets of clusters could be presented in the form of a map with contour lines similar to those used by surveyors.
3. Centroid factor analysis followed by application of the simple structure concept of Thurstone (22) as a basic aid for the validation of items and the interpretation of primary scores.
4. Fundamental principles of matrix theory as a means of maximizing validity and reliability with a relatively small number of items.

Some psychologists, at the time this program was initiated, thought there were probably no more than 20 primary personality traits. Our 20 year study, however, yielded 128 thoroughly validated primary scales and two scales of the second order.

### *1. Preliminary Treatment of Item Responses*

A systematic approach involving two methods, namely, content analysis and cluster analysis, was always utilized in the early phases of this investiga-



tion. The results of content analysis of the items was compared with that of cluster analysis of the responses to the same items. Because of the laborious task of computing a large number of intercorrelations, the inventory was divided into sets of 50 items. The results of the two methods were compared for each item set. After this series of operations was completed, 30 to 50 item sets selected from the entire inventory were again subjected to cluster analysis.

In the early phase of this study, items were not accepted for further consideration unless the results of the two methods were in substantial agreement. Later, rigid adherence to this approach was abandoned for the following reasons. First, the two methods often generated conflicting results for items of low face validity. Consequently, cluster analysis or factorial methods seemed to be the only safe procedures for classifying such items. Second, some of the most contradictory data seemed to be related to semantics, a topic which will now be critically examined.

## 2. *Serious Contradictions between the Results of Content Analysis and Cluster Analysis*

The type of contradiction to be discussed here is serious because of its basic importance for the construction of valid personality scales. It was disturbing to find that items which seemed to be opposite in meaning proved to be statistically independent when subjected to cluster analysis. In other words, they tended to be uncorrelated instead of yielding negative correlations. In fact, this difference between semantic opposites was so clear-cut, that it was readily revealed by merely inspecting the correlation coefficients. When such antonymous items were classified into scoring categories by means of cluster analysis, there also tended to be no significant correlations between the scores. Let us examine a small set of tetrachoric relationships for 200 undergraduate women (Table 1).

TABLE 1  
SHOWING THE STATISTICAL INDEPENDENCE OF TRAITS WHICH ARE POPULARLY CONSIDERED AS ANTONYMOUS

Scoring category	Group I		Group II	
	1	2	3	4
I 1. Buoyancy (tendency to be happy)	1.00			
2. Dynamism (tendency to feel energetic)	.58	1.00		
II 3. Depression (tendency to feel unhappy)	-.06	-.10	1.00	
4. Lethargy (tendency to feel drowsy or tired)	.00	-.11	.61	1.00

In this table, there are no significant correlations between semantic opposites, such as buoyancy and depression. The terms, buoyancy and depression, as defined here, are antonyms. As antonymous categories, there should be an extremely significant negative correlation between the two, thus indicating that they are at opposite poles of a single trait continuum, thus supporting the bipolar concept of personality traits.

Our data, on the contrary, show that there is no significant relationship between buoyancy and depression, thus indicating that we are dealing with *two* traits instead of a *single* bipolar dimension. Our quantitative analysis showed that this was true for all semantic opposites which refer to traits in the emotional and motivational areas of behavior. One interesting consequence of this unipolar concept is that the number of primary scores will be twice that for a corresponding set of bipolar traits.

A similar result emerged when inter-item correlations were factor analyzed. One published report lends indirect support to this point of view. Thurstone (23) found no evidence of bipolarity when he factor analyzed the 13 scores of the Guilford series of personality measures. "We started this analysis," he states, "with the expectation of finding bipolar factors for all or most of these factors, but the result revealed all of them to be positive."

### 3. *Bipolar Versus Unipolar Concepts of Personality Organization*

Inspection of Table 1 suggests that semantically opposite scores tend to fall into separate groups of related scales. Here the within-group covariance is relatively high but the between-group covariance is extremely low. Buoyancy, for example, correlates significantly with Dynamism and the same is true for Depression and Lethargy. On the other hand, the lower left quadrant of the table reveals the fact that the inter-group coefficients are statistically insignificant.

Data, such as these, are apt to be quite troublesome to individuals who have been conditioned throughout life to the semantics of our modern languages. In 1950, in a report to the New York Academy of Sciences, the writer (11) attempted to explain this apparent paradox with the following statement: "Such results may seem confusing if we fail to keep in mind the fact that *behavioral tendencies* rather than *behavioral effects* are uncorrelated. For example, possessing both buoyancy and depression to a high degree does not mean that an individual must be unhappy and depressed at the same time."

Examination of Figure 1 suggests that affective and conative primary traits tend to cluster into two statistically independent groups instead of appearing at opposite poles of a single factorial axis. This chart is based on a



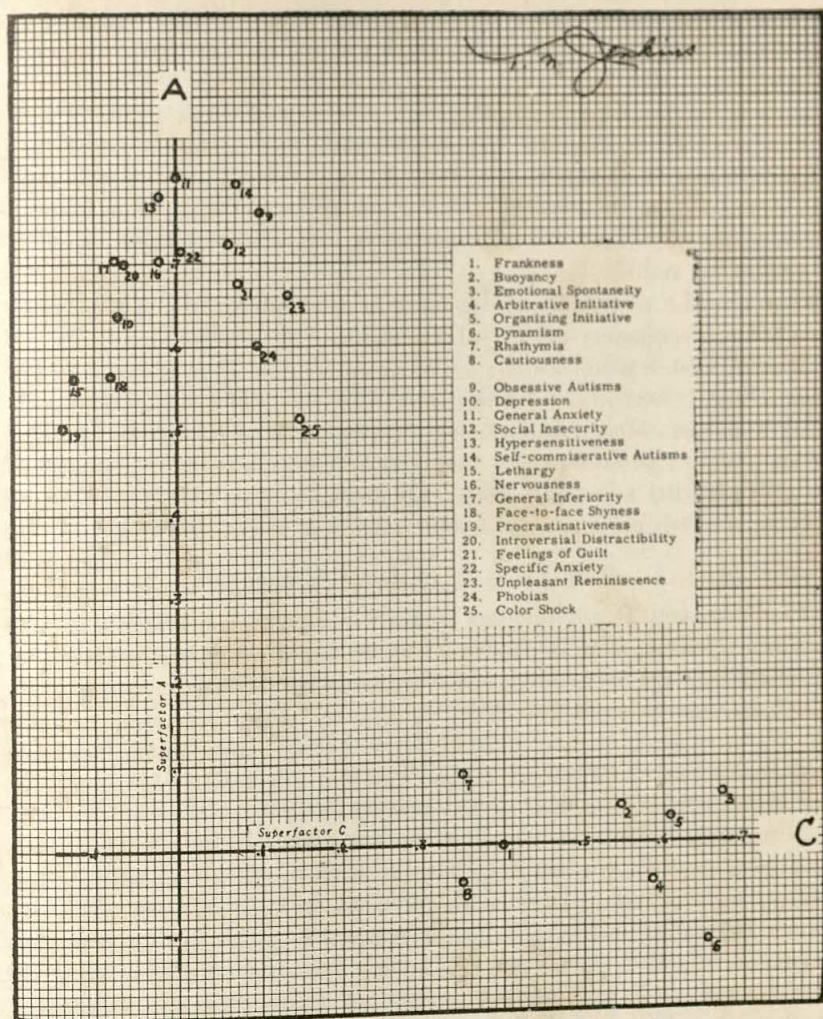


FIGURE 1

Graphical representation of relationships between primary test scores and superfactors. The numbers within the gaps of the heavy lines are the factor loadings (correlations with the superfactor).

centroid factor analysis of the primary scores for 500 military cadets. Since the scores are primary, the factors which emerge are secondary (22). In this chart, these factors of the second order are called superfactors.

It is obvious that these results are important for the construction of valid tests of personality. Our studies, moreover, serve to emphasize the necessity of rigorously checking all items in a scoring category for factorial relevance. They also add another distinction between personality traits and attitudes, the traits being unipolar while attitudes are bipolar. Finally, these findings contraindicate prevalent and past tendencies to construct bipolar measures of personality traits.

Cattell's (7) comprehensive summary of trait studies shows that practically all investigators accept, either implicitly or explicitly, the bipolar concept of trait organization. It is worthy of note that this view is accepted implicitly by laymen, a fact which is especially important for semantics. The English language semantically assumes the acceptance of the bipolar concept, exemplified by the definition of trait terms in our dictionaries. It appears that psychologists and psychiatrists exhibit an habitual tendency to accept uncritically this universal lay notion regarding personality organization.

### C. THE PROBLEM OF SCORE VALIDATION

This 720 item global inventory yields 139 scores, thus generating some serious problems of score validation. There are 128 primary factor scales, two factor scales of the second order and nine non-factorial scales.

#### 1. *Validation of the Primary Scales*

An extensive test battery of over 100 scoring categories serves to remind one of the need for making a highly parsimonious use of items. Otherwise the list would become much too long for any practical use. On the basis of traditional methods, one would be likely to select a large number of items to estimate a given factor. To do this, one would select those items which correlate significantly with the factor. A large set of items might result in a high level of reliability without producing a high level of validity. Instead of the usual traditional methods, a systematic procedure was needed for maximizing validity with a relatively small number of items.

(a) *Maximizing validity through the aid of modern theories of matrices.* One way to utilize items in a highly parsimonious manner is to select those that will yield intercorrelations constituting what mathematicians call a unifactor matrix. The best estimate of a primary trait can be achieved when the rank of the correlation table is *one*. A matrix of rank one will yield



only one unique factor. This tends to minimize the effects of contamination by irrelevant factors. Important properties of this type of correlation matrix were discovered by Spearman (20) without the use of matrix algebra. In fact, a unifactor matrix is often referred to as a Spearman hierarchy. Burt (5) has directed attention to the importance of the unit hierarchy. Later, Thurstone (22) suggested that the ideal distribution of irrelevant factors is one in which no two items contain the same irrelevant factor variance.

It is difficult to select a set of items which will completely satisfy conditions for a unifactor matrix. The alternative is to select sets which will approximate this ideal as closely as possible. If the matrix of intercorrelations approximates a Spearman hierarchy, high score validities can be achieved with a relatively small number of items. The need for parsimony in developing this extensive battery of personality tests forced us to spend several years in search of item sets which would closely approximate the unifactor ideal.

(b) *What is a primary scale?* There has been some confusion regarding the meaning of certain factorial terms. We might define a primary scale as one based on a factorial criterion. This statement, however, hardly clears the air. We need a definition of a primary factor. Thurstone (22, p. 348), after describing a primary vector as a unit vector defined by a coordinate axis, states that a "trait which corresponds to a primary vector will be called a *primary trait* or a *primary factor*." This definition, which is difficult for psychologists with little mathematical training to understand, fails to differentiate *primary* factors from factors of the *second order*. Those who often utilize Thurstone's methods tend to refer to primaries as the factors which emerge as the outcome of the *initial* analysis of a set of items or behavioral variables (6, 10, 21, 22). This writer accepts this practical notion of primaries. In like manner, *secondary* factors would be those resulting from a *second* analysis, that is, the factor analysis of the intercorrelations of the primary factors.

(c) *Discriminating power and reliability.* Computation of the discriminating power and reliability of items can be a valuable aid in selecting items which will increase score validity. Every item utilized in this series of scales has been checked for response variance and retest reliability. An attempt was made to select items of high discriminating power which at the same time would tend to satisfy the conditions for a unifactor matrix. Under these circumstances, we were not uniformly successful. The broad, six-point response scale facilitated this task. Moreover, the broad response scale simplifies the task of selecting satisfactory cutting points for scoring.

When the retest reliability of an item was low, it was rejected. This was

done on the assumption that the retest reliability of items would be reflected in the retest reliabilities of the scores. The unreliability of most of the items in this test battery was also separated into *error* variance and *personal* variance (quotidian variation). By personal variance, we mean a tendency to vary from time to time which is a trait characteristic of the individual. Thus, it should be clearly distinguished from error variance. This separation was done by utilizing a procedure suggested by Fagin (10) for breaking up the unreliability of items into error variance and personal variance. Woodrow (10) in his discussion of reaction time was the first to report this type of variance, which he named quotidian variation.

(d) *Final verification of scale validities.* In determining the rank and validity of a scoring category, inter-item correlations were always subjected to centroid factor analysis followed by graphic rotation to an orthogonal simple structure (22). Recently these results were verified through the use of electronic computers. By means of the Univac installation at New York University, the matrices were factored by the method of *principle axes* (22). Later, at the Watson Laboratories an IBM 650 was utilized to factor the scoring matrices by a centroid program developed by Jonathan Robbin (4). Rotation to an orthogonal simple structure was accomplished electronically by means of the quartimax method of Neuhaus and Wrigley (17).

## 2. Construction and Validation of the Superfactor Scales

Returning to Figure 1, it will be noted that there are two clear-cut and statistically independent clusters of test scores associated with superfactors *A* and *C*. These outstanding superfactors were first reported by the writer in 1950 (11). As a result of theoretical considerations, superfactor *A* is now referred to as *adrenergic* and superfactor *C* as *cholinergic* (11, 12).

(a) *Construction of the scales.* Figure 1 shows that indirect individual estimates of adrenergic and cholinergic could be estimated readily by introducing the primary scores into regression equations. A direct procedure, utilizing the actual responses to test items, would be more convenient. The direct approach, moreover, would probably generate test scores which are more valid than those obtained through regression.

The construction of these measures was a very laborious task. Using the primary scales, regression scores were computed for each subject in four diverse samples (undergraduate men and women, basic airmen and retail saleswomen), each comprising about 200 to 300 individuals. Then the regression estimates were correlated with each response to the 720 item inventory for each sample. Because of error variance, not only in the item



responses but also in the estimated scores, the coefficients were apt to vary considerably in magnitude and sign. An item was not accepted for a superfactor scoring category unless the item correlations were substantially the same for the four populations.

It is important to note that these two scales were statistically independent. The Pearson correlations between cholinergy and adrenergery are  $-.07$ ,  $-.04$ ,  $-.03$  and  $-.06$ , respectively, for undergraduate men, undergraduate women, basic airmen and retail saleswomen.

(b) *Final validation of the superfactor scales.* The validity coefficients of these scores are simply the correlations between the scores and the superfactors (14). Four populations (323 undergraduate men, 200 basic airmen, 351 undergraduate women and 324 retail saleswomen) were utilized to determine score validity. All of the test scores, comprising these two scales and 128 primary scores, were factor analyzed by an IBM 7090 electronic computer with a special program devised by Jonathan Robbin. The normal varimax method of Kaiser (15) was then used to rotate the superfactorial matrices to an orthogonal simple structure.

The resulting validities (superfactor loadings) were satisfactory and very uniform for the four populations. The mean validity coefficients for *Adrenergery* and *Cholinergy* were .88 and .86 respectively.

### 3. The Non-factorial Scales

Several studies were executed to develop nine non-factorial scales. These comprised measures of test distortion, test consistency, conservatism, self-acceptance, adjustmental discrepancy and authoritarianism.

(a) *The test objectivity scale.* This is probably the most important of the distortion tests. This objectivity scale is especially useful for an over-all evaluation of individual test results. It was designed to ascertain how realistically the examinee answered the items. It provides a measure of the examinee's success in presenting a realistic personal portrait of himself. A low objectivity score, for instance, usually indicates an automatic or non-deliberate tendency to create a favorable impression.

This scale evolved over a period of several years. The present discussion will be limited to a brief outline of a few phases of this test development.

(1.) In 1946 the writer interviewed about 200 students to find two extreme groups, namely, those who appeared to have very realistic conceptions about themselves and those who seemed to be highly unrealistic. These students had taken the experimental battery of 1204 items. Significant differences

between these two criteria were revealed for 187 items. From this group of items, 63 were selected for a preliminary scale.

Later an attempt was made to construct separate measures of deliberate and non-deliberate distortion. Some of the operations utilized to reveal automatic tendencies to give unrealistic responses will now be described briefly.

(2.) Forty-two graduate students indicated on a six-point scale what they believed to be unrealistic responses to each of the 187 items.

(3.) Next, regular test results for 186 students were compared with those involving *deliberate* attempts to falsify in an experimental inventory of 300 items. The students at one time attempted to make a *convincingly favorable* impression while at another time they tried to make a *convincingly unfavorable* impression. This operation generated the following results: (a) significant differences between attempts to make favorable and unfavorable impressions; (b) how these differential effects disagreed with those of the two criterion groups in the 1946 study; and (c) how these effects differed from the 42 graduate judges.

(4.) A systematic study of each cell in the six-point response scale was executed for this group of students. This operation produced two valuable results. First, we were able to determine the relative validity of various cell responses as indicators of test objectivity. Second, on the basis of retest data, it was found that many item responses were too unstable to include in an objectivity scale.

Scales of this type are difficult to validate in quantitative terms. As the result of student interviews, the writer concluded that the scale tends to furnish a good estimate of what it is supposed to measure.

The nature of test objectivity as a subjective test attitude is revealed in part by the correlations between this scale and various primary traits. A few of these relationships are shown in Table 2. The correlations are positive for socially disapproved traits and negative for approved ones. Moreover, the coefficients tend to be larger for disapproved than for approved traits, a fact which may be characteristic of our western culture.

(b) *The other distortion scales.* A systematic attempt was also made to select items to estimate deliberate distortion and self-depreciation. The data obtained in constructing the objectivity scale were utilized in the development of the other distortion categories. In addition, the 720 item inventory was administered to a new sample with the three different sets of instructions previously mentioned. Ten judges also indicated which responses to this inventory were likely to be unrealistic, and hence unconvincing. These



judges comprised one clinical psychologist, two academic psychologists, two IBM operators and five students. These judgments and the item distributions were used to construct preliminary scales of sincerity and self-depreciation.

TABLE 2  
SHOWING POSITIVE CORRELATIONS BETWEEN TEST OBJECTIVITY AND SOCIALLY DISAPPROVED TRAITS AND NEGATIVE RELATIONSHIPS BETWEEN APPROVED TRAITS

Scoring category	Male students (N = 323)	Air trainees (N = 200)	Female students (N = 351)	Retail saleswomen (N = 324)
1 Irritability	.26	.36	.21 <sub>s</sub>	.38
41 Depression	.24	.33	.21	.26
54 Distrustfulness	.28	.46	.31	.42
62 Psychosomaticism				
68 Undesirable compulsiveness	.40	.52	.32	.51
112 Unpleasant reminiscence	.36	.47	.27	.43
73 Negative algolagnia	.40	.43	.38	.45
114 Fearfulness	.48	.62	.43	.48
136 Adrenergy	.35	.51	.32	.57
19 Novelty-loving	— .24	— .27	— .24	— .20
21 Vocational assurance	— .16	— .36	— .27	— .29
45 General morale	— .18	— .30	— .17	— .14
49 Persistence	— .19	— .36	— .29	— .28
71 Intellectual adequacy	— .25	— .35	— .20	— .27
76 Motor tempo	— .27	— .31	— .18	— .23
90 Emotional control	— .16	— .31	— .13	— .16
135 Cholinergy	— .16	— .22	— .17	— .16

While developing a preliminary scale for test sincerity, it was noted that examinees behaved differently to culturally acceptable statements than to unacceptable ones. For instance, an individual might falsify in reacting to "I like teamwork" and be sincere to "I like to torment others." Hence, two sincerity scales were constructed. *Sincerity I*, containing 21 items, comprises only statements, the content of which is usually *approved* in our culture. *Sincerity II*, containing 35 items, comprises statements which involve culturally disapproved conduct. Low response loadings to such items would correlate *positively* with cultural approval. For instance, a response of *never* to "I like to torment others" could be a culturally approved response, although clinically it might be considered a little abnormal.

While constructing the self-depreciation scale some interesting but serious difficulties arose. Among these were peculiarities of the item response distributions when students were requested to answer the items so as to make a convincingly *unfavorable* general impression. The modal response for some items was in the central portion of the six-point response scale. As stated heretofore, the six possible responses to an item were as follows: *always*,

*almost always, often, sometimes, hardly ever, and never.* The modal response for other items tended to appear at either end of the response scale. But some of the response distributions were *bimodal*, that is, two different portions of the scale were considered as culturally unfavorable. Let us call the unimodal items *univalent* and the bimodal items *bivalent*. Two self-depreciation scales were produced, namely, *univalent* and *bivalent* scales. Some questionable assumptions have to be made to score the bivalent items. Hence, only the univalent scale has been retained for the A-42 revision of this test battery. This self-depreciation scale contains 37 items.

(c) *The test consistency scale.* This scale was designed to indicate degree of consistent performance while taking the test. It was not supposed to serve as a predictor of consistency in taking tests nor consistency when taking a test on *another* occasion. It is strictly a measure of consistency for a *given* occasion. A low score would cast some doubt on the usefulness of an individual's test results.

This score is based on the consistency of reaction to items of the same content. Ten items in the first portion of the test are repeated toward the end of the test. These items were selected so as to balance diverse forms of response distributions, such as distributions skewed to the left and to the right, and others which were symmetrical. The sum of the response differences between the first set of items and the identical repeat items constitutes a raw inconsistency score. The consistency score is obtained by subtracting this sum from a *maximum possible* score of 50.

The correlations between consistency and the factorial personality scores strongly suggest that test consistency is not an estimate of a unique personality trait. These relationships are not only small but very few of them are significant statistically. Ninety-two per cent of these correlations falls within the interval  $-.10$  to  $.10$ , while ninety-nine per cent falls within  $-.15$  and  $.15$ .

The relationships between consistency and distortion on the contrary, are all statistically significant. The coefficients are small but they all tend to indicate a negative relationship between consistency and distortion. Highly consistent individuals, for instance, are likely to be more objective than inconsistent ones. The correlations between consistency and the various distortion scales are presented in Table 3.

(d) *The conservatism scale.* This scoring category was designed to furnish an over-all estimate of the following behavioral trends: (a) tendency to *conform* to conventional, middle-class values; (b) tendency to *conserve* conventional, middle-class values and to *actively* oppose changes in the customary way of doing things. As indicated previously, Lentz's C-R Opinion-



aire was utilized in our early research work as a criterion of conservatism (16). The combined score for Forms J and K was used for this purpose in 1946.

TABLE 3  
SHOWING THE PEARSON COEFFICIENTS BETWEEN TEST CONSISTENCY AND TEST  
DISTORTION FOR FOUR POPULATIONS

Distortion category	Under-graduate men	Air Force trainees	Under-graduate women	Retail sales-women
Test objectivity	.16	.23	.21	.15
Sincerity I	.16	.19	.16	.26
Sincerity II	.14	.22	.26	.16
Self-depreciation	-.14	-.22	-.15	-.29

After a lapse of a dozen years, many of the items in the C-R Opinionaire were no longer timely. Hence, in 1958, 60 items were selected from Forms J and K which were not only maximally discriminating, but which correlated highest with the combined test score. This 60 item list was presented to approximately 300 students. These scores were utilized to select those items of the global personality battery (A 42) which would furnish the best estimate of conservatism, as defined in the previous paragraph.

(e) *The measurement of self-acceptance and adjustmental discrepancy.* Clinical writers have long emphasized the adjustmental importance of a patient's concept of himself. Self-esteem seems to be the essential factor implied by writers about the self-concept (19). Several investigators, however, have shown that the self-concept is factorially complex. In fact, one investigator has enumerated nine clinical criteria of self-acceptance (2).

Bills (3) and his co-workers developed a fairly simple procedure for sampling the concept of the real self, self-acceptance, and the concept of the ideal self. A modification of this procedure was utilized to obtain self-acceptance and adjustmental discrepancy scores for approximately 300 students. The 49 adjectives selected by Bills and his co-workers were repeated on three different pages. On the first page, the examinee was instructed to show how often he is like the quality indicated by the adjective. The following frequency response scale was used: *always, almost always, often, sometimes, hardly ever, and never.*

On the second page, the examinee was requested to designate how well he liked what he is. For instance, does he like or dislike being "cruel"? The following response scale was utilized: *like tremendously, definitely like, like slightly, dislike slightly, definitely dislike, and dislike tremendously.* This page furnished the data for a self-acceptance score.

On the third page, the examinee responded to the same frequency scale used on the first page. But this time he responded by showing how often he *would like* to be like the quality indicated by the adjective. For instance, how often would he like to be "annoying"? The differences in response to identical items on page one and page three were summed. This sum was the adjustmental discrepancy score, that is, the difference between the examinee's concept of real self and the concept of an ideal self. It is an estimate of the difference between what the individual thinks he is and what he would like to be.

These scores were correlated with the 720 items of the global personality battery. In this way, items were selected for obtaining direct measures of self-acceptance and adjustmental discrepancy. The Self-acceptance scale contains 27 items while the Adjustmental Discrepancy contains 30 items.

(f) *The Authoritarian scale.* In 1950, Adorno (1) and his co-workers, in a comprehensive study of authoritarian personalities, published a tentative test of authoritarianism. This scale has often been referred to as the F-test on the assumption that it was a measure of fascism. These items with a six-point response scale were administered to about 300 students. The scores were correlated with the 720 items of the personality test. Thus, the F-test was used as a criterion for the selection of personality items for an authoritarian scale. This authoritarian scale, in its final form for A-42, contains 46 items.

A few words of caution should be offered concerning the use of this scale. It was generally believed that extreme conservatives were likely to be authoritarians. Careful graduate studies at New York University, however, have shown that extreme radicals are also likely to return high authoritarian scores. The extreme radical strongly rejects conventional, middle-class values. The extreme conservative, on the other hand, would strongly oppose those who fail to conform to conventional, middle-class values. For these reasons, the conservatism scale should always be considered for an accurate appraisal of the authoritarian score. In other words, there are conservative authoritarians and radical authoritarians.

#### D. RELIABILITY

The reliability of this test battery was ascertained by the retest method. This operation was carried out for 139 variables. Among these were 130 factorial scoring categories and 9 non-factorial categories.

The normative retest reliabilities of these tests were determined for undergraduates in Washington Square College of New York University. These samples comprised 117 women and 157 men. The retest interval was five weeks. The retest reliabilities for the 139 categories are presented in Table 4.



TABLE 4  
RETEST RELIABILITY COEFFICIENTS FOR UNDERGRADUATE MEN (N = 157) AND UNDERGRADUATE WOMEN (N = 117)

Since the items were factored, the usual measures of internal consistency, often referred to as reliability, were not utilized. Such measures of internal consistency are not appropriate for tests based on unifactor validities. Factor analysis not only reveals the consistency of a set of items but also furnishes estimates of both relevant and irrelevant factor loadings. Several writers (24, 25) have challenged some of the common assumptions utilized in ascertaining internal consistency (reliability). Wherry and Gaylord (25) have shown that many procedures for determining reliability produce misleading results because of failure to take account of factorial complexity. In fact, they conclude that a complete factor analysis is the proper approach to the problem of reliability.

At this point, I would like to direct attention to a fact which is usually overlooked by those who are accustomed to traditional methods of test construction. This important fact is the relationship between reliability and factorial complexity (25). Level of reliability of a score tends to be *inversely* related to the factorial complexity of a set of items. A large number of factors, in other words, suggests a large number of primary reactions for generating the test score, thus lowering reliability. It follows, therefore, that the *ideal set of items for maximizing factorial validity is also the ideal for obtaining high retest reliability with a small number of items.*

This principle is well exemplified by the low reliabilities (Table 4) for test consistency (score no. 131), since consistency can vary in many different ways, such as the following:

1. A deliberate attempt to make a favorable impression.
2. Failure to take the test seriously.
3. Unrealistic conceptions about oneself. There is a significant relationship between Test Objectivity and Consistency (Table 3) thus suggesting that a lack of realism about oneself may tend to generate variable responses to test items.
4. Uncertainty about oneself.
5. A shift in test attitudes while taking the test.
6. Low self-acceptance, as revealed by the self-acceptance score.
7. An habitual tendency to respond to situations in an unstable manner.
8. An habitual tendency to be careless (see scores for personal recklessness and caution).
9. Misunderstanding the test items due to language difficulties.



## E. DESCRIPTIONS OF THE SCORES

Brief descriptions are convenient but in some cases a dictionary definition of a one-word title fails to convey the basic meaning of the scale. There are also many primary scales, such as *vestibular empathy*, which are new in psychology. Therefore, in the following list of 139 scales, the name of each scale will be accompanied by a concise description.

1. *Irritability*—Tendency to become annoyed or upset when one feels thwarted or frustrated by people or conditions.
2. *Face-to-face initiative*—Tendency to take the initiative in meeting strangers and putting them at ease.
3. *Mutual secretiveness*—Tendency to share personal secrets with only close friends.
4. *Practicality*—Tendency to think and to cope with the environment in relation to practical or utilitarian needs.
5. *Cautiousness*—Tendency to be prudently attentive to dangers or risks in doing work or choosing friends.
6. *Dependent initiative*—Tendency to actively seek help, or to make requests for help or sympathy.
7. *Independence (aid-avoidance)*—Tendency to actively avoid receiving help or advice from others.
8. *Vestibular empathy*—Tendency to experience equilibril sensations in response to moving objects or living beings which are merely seen or heard, instead of depending strictly on the position or movement of the observer's body. Sometimes described by subjects with high scores as a feeling of "gliding or floating."
9. *Desirable compulsiveness*—Tendency to feel impelled to do socially approved tasks in a certain way, and often to carry them out to excess, as for example cleaning more than is necessary or practical.
10. *Frankness*—Tendency to be completely frank and outspoken in stating opinions about issues, things, and other people, even when it is not tactful or politic to do so.
11. *Buoyancy*—A resilient tendency to be happy.
12. *Punctuality*—Tendency toward promptness or concern about timeliness in meeting schedules and commitments as opposed to tardiness or dilatoriness.
13. *Cultural Non-conformity*—Tendency to be unconventional, that is, to feel free to do things differently from others.
14. *Emotional reticence*—Tendency to keep one's feelings to oneself.
15. *Social assurance*—Tendency to feel that one is liked and accepted by others.
16. *Personal recklessness*—Tendency to indulge in physically dangerous activities.
17. *Persuasiveness*—Tendency to influence others to agree with one's own way of thinking or of doing.
18. *Pleasant reminiscence*—Tendency to daydream about pleasing experiences in the past.
19. *Novelty-loving*—Tendency to like situations involving new decisions, plans, goals or new ways of doing things in contradistinction to mere variety without novelty.
20. *Remedial aggressiveness*—Tendency to request others to remedy frustrating or unpleasant situations.

21. *Vocational assurance*—Tendency to feel capable of achieving the financial or occupational rewards which one needs or desires.
22. *Building up others (ego-building)*—Tendency to try to increase a feeling of self-importance (ego-strength) in others.
23. *Cooperativeness*—Tendency to identify oneself with and work with others to achieve common goals.
24. *Vocational self-sufficiency*—Tendency to feel no need for help or advice from others in carrying out important tasks.
25. *Vindictiveness*—Tendency to be revengeful or retaliatory.
26. *Service-mindedness*—Tendency to help others with their difficulties in everyday activities.
27. *Sympathy*—Tendency to experience the feelings or emotions of others, whether joyful or sorrowful.
28. *Ambitiousness*—Tendency to desire and to strive for personal preferment or advancement in the sense of seeking marks of success or prestige, such as fame, honor, money, and influence.
29. *Disputatiousness*—Tendency to engage in debate, argumentation, or controversial discussion as a mode of intellectual self-expression. Usually provocative controversial behavior.
30. *Hypercriticalness*—Tendency toward noticing and pointing out the faults, mistakes or shortcomings of others as well as a tendency to direct attention to and over-emphasize these faults.
31. *Piety*—Tendency to be reverential or devout and faithful in the performance of moral obligations.
32. *Weepiness*—Tendency to cry or weep easily.
33. *Emotional spontaneity*—Tendency to express oneself in an uninhibited, natural and enthusiastic manner.
34. *Superiority feelings*—Tendency to be aware of one's own superiority and when marked, often coupled with a tendency to be unrealistic about it.
35. *Indecisiveness*—Tendency to have difficulty in making decisions, thus often resulting in a lack of timeliness in making up one's mind.
36. *Status seeking*—Tendency, sometimes deceptively, to impress others with respect to one's own importance, worth, or status.
37. *Ruminative autisms*—Tendency toward uncontrolled repetition of relatively unimportant thoughts.
38. *Uncooperativeness*—Tendency to be non-cooperative and actively resistant to teamwork or cooperative activity.
39. *Cathetic obedience*—Liking to obey, involving a tendency to seek situations in which obedience is a highly acceptable form of conduct. A person may be submissive, yet not like to obey superiors (cf 81).
40. *Conscientiousness*—Feeling of *oughtness* or moral obligation. Tendency to be influenced by a sense of right or wrong together with a feeling of obligation to do or be that which is believed to be right (cf 69).
41. *Depression*—Tendency to feel unhappy, dejected, low-spirited or disheartened.
42. *General anxiety*—Tendency to worry or feel anxious without reference to specific situations. Sometimes referred to as free-floating anxiety.
43. *Social insecurity*—Tendency to feel that one is not liked, loved, approved, or accepted by others.
44. *Sexual inferiority*—Feeling inadequate in relation to the opposite sex.
45. *General morale*—A general tendency to feel optimistic, hopeful, or satisfied with respect to both the world at large and one's present and future status in it.



46. *Hypersensitiveness*—Feelings tend to be easily hurt.
47. *Self-commiserative autisms*—Tendency to indulge in self-pity.
48. *Cycloidism*—Tendency to oscillate from sadness or lethargy to happiness or liveliness, or the reverse.
49. *Persistence*—Tendency to continue against opposing forces, either outside or within oneself.
50. *Lethargy*—Tendency to be unenergetic, drowsy, or tired.
51. *Nervousness*—Tendency to feel jumpy, tense, restless, or unable to relax, sometimes manifested by fidgeting and exhibiting nervous habits known as tics.
52. *Selfishness*—Tendency to take advantage of people or situations without regard to the happiness, welfare, or rights of others.
53. *Unwilling cooperativeness*—Tendency to be overtly cooperative without feeling cooperative.
54. *Distrustfulness*—Tendency to be suspicious of others.
55. *Overt self-depreciation*—A conscious tendency to present an unfavorable view of oneself.
56. *Need for dependence*—Liking to be dependent on, or "mothered," by others.
57. *Self-algolagnia*—Tendency to enjoy hurting or punishing oneself.
58. *Procrastination*—Tendency toward inexcusable delay in doing what one wishes to do.
59. *Positive algolagnia*—Tendency to derive satisfaction from the unhappiness, discomfort, or embarrassment of others, and to enjoy making others feel uncomfortable.
60. *General feeling of inferiority*—A general feeling of inadequacy, or over-all weakness of personality. A free-floating feeling of inferiority.
61. *Interest in physical combat*—Examples: boxing contests, wrestling, bull fights, dog fights, etc.
62. *Psychosomaticism*—Tendency to exhibit symptoms of psychosomatic disorders, that is, bodily diseases which are supposed to be related to psychic disturbances. Example: anxiety and gastric ulcers.
63. *Arbitrative initiative*—Tendency to take the lead in settling disputes or disagreements in everyday situations.
64. *Organizing initiative*—Tendency to take the lead in organizing business or ordinary everyday (non-business) activities.
65. *Intuitiveness*—Tendency to depend and act on the unreasoned solution of problems. Such solutions are popularly referred to as hunches.
66. *Face-to-face shyness*—Tendency to feel bashful, self-conscious, and embarrassed in dealing with authorities, important people, and strangers.
67. *Grandiose autisms*—Tendency to be involved continually with thoughts of personal grandeur, such as, daydreaming that one is a great hero, leader, scientist, and so on.
68. *Undesirable compulsiveness*—Feeling a strong need to do something which is socially unacceptable.
69. *Seriousness*—Tendency to assume an attitude of earnestness or to exhibit feelings of personal responsibility toward one's work or environment. A person may be serious without being conscientious (cf 40).
70. *Homotropism*—Tendency to identify more or less completely with members of one's own sex, but not necessarily in a homosexual manner.
71. *Intellectual adequacy*—Feeling that one's abilities are adequate to cope successfully with one's own problems.

72. *Rebelliousness*—Resistance to domination. Tendency to actively resist ascendant or dominant behavior on the part of others. Implies open defiance of authority.
73. *Negative algolagnia*—Tendency to enjoy being hurt by others, or to derive satisfaction from mild pain.
74. *Social initiative*—Tendency to take the lead in social gatherings.
75. *Gluttony*—Liking to eat to excess.
76. *Motor tempo*—The habitual rate or vigor of movement which one tends to exhibit in everyday or routine activities.
77. *Gregariousness*—Tendency to need and to seek to be with other people.
78. *Sexual romanticism*—Tendency to idealize relationships between the sexes. May be unrealistic.
79. *Introspectiveness*—Tendency to analyze, and be preoccupied with, one's own thoughts or feelings.
80. *Domineeringness*—Tendency to be "bossy", imperious, or over-bearing in dealing with other people.
81. *Submissiveness*—Tendency to comply with what others expect one to do and to submit to domination rather than to resist it (cf 39).
82. *Impulsiveness*—Tendency to act precipitately or to make hasty decisions without careful consideration or deliberation.
83. *Seclusiveness*—Tendency to be, and to prefer to be, alone.
84. *Vocational insecurity*—Feeling that one is likely to be unsuccessful in achieving the financial or occupational rewards which one needs or desires.
85. *Rhathymic exhibitionism*—Tendency to act in a carefree manner in the presence of others, that is, to dramatize carefreeness. Exhibited in various ways, such as playing practical jokes. A person may put on an exhibition of carefreeness, yet not feel carefree (cf 102).
86. *Mucodermal orexis*—Tendency to enjoy and to seek stimulation of the cutaneous and membranous surfaces of the body.
87. *Dynamism*—Tendency to feel lively or energetic, with a sense of vigorous physical well-being.
88. *Tidiness*—Cleanliness and neatness as opposed to slovenly behavior.
89. *Kinesthetic empathy*—Tendency to have kinesthetic experiences (muscular or movement sensations) in response to activities which are seen or heard.
90. *Emotional control*—Tendency to inhibit or restrain socially disapproved emotional reactions, such as controlling one's temper in disagreeable situations or remaining calm when others become upset.
91. *Variety-loving*—Liking changes from time to time with respect to situations, things, or jobs, which one is accustomed to. A person may like variety, yet dislike the new (cf 19).
92. *Vistorexis*—Tendency to enjoy and to seek scenery, distant views, vistas, and so on.
93. *Introversional distractibility*—Tendency to be distracted from intended goals by disturbances within one's own body, such as ideas and internal sensory experiences.
94. *Talkative gregariousness*—Preferring to talk rather than listen, when with groups of people.
95. *Forgiveness*—Tendency to forgive and to overlook the wrongs or mistakes of others toward oneself.
96. *Extraversional distractibility*—Tendency to be distracted from the achievement of intended goals by irrelevant external conditions or stimuli.



97. *Spatial disorientation*—Tendency to feel uncertain about directions or positions in space.
98. *Defensive projection*—Tendency to project one's own faults on other people. Example: if one is cowardly, he may tend to assume that other people are cowards.
99. *Feelings of degradation*—Tendency to feel guilty or ashamed of oneself.
100. *Situational anxiety*—Tendency to worry about specific, probable, or possible dangers, risks, or misfortunes.
101. *Carefree deception*—Overtly carefree but covertly serious, cautious, or both.
102. *True rathymia*—Tendency to feel carefree, but not necessarily in an irresponsible way. Should be clearly distinguished from rathymic exhibitionism (cf 85).
103. *Control of disagreeable feelings*—Tendency to keep disagreeable feelings, such as disapproval or dissatisfaction, to oneself.
104. *Ambitious agreeableness*—Tendency to act friendly or agreeable in order to achieve or succeed.
105. *Hoarding*—A compulsive tendency to save, often directed to apparently unimportant things on the assumption that they may be useful in the future.
106. *Intellectual inferiority*—Tendency to feel that one is slow to learn or to grasp the meaning of new ideas.
107. *Public shyness*—Tendency to feel uncomfortable in the presence of groups of people outside the family circle.
108. *Theoretical thinking*—Tendency to be interested in, and to solve problems, involving theories, or abstract principles, as contrasted with solving problems to meet immediate practical objectives.
109. *Anti-social recklessness*—Tendency to take chances in breaking rules, such as, traffic regulations while driving, and smoking where forbidden.
110. *Anti-disputatiousness*—Tendency to avoid arguments or provocative discourse.
111. *Defensive secretiveness*—Tendency to conceal facts about one's personal weaknesses.
112. *Unpleasant reminiscence*—Tendency to be preoccupied with unpleasant memories.
113. *Conservatism*—Tendency to approve commonly accepted or standard customs and practices as well as a reluctance to depart from them.
114. *Fearfulness*—Tendency to be fearful in the presence of real objects or situations. Habitual fears are sometimes referred to as phobias. Does not comprise conjectural or obsessional fears (worries), such as, specific anxieties (cf 100).
115. *Color shock*—Tendency to get upset by bright colors or outstanding grays (black or white).
116. *Naso-buccal epicurism*—Tendency to be preoccupied with or interested in chemoreceptive stimulation, especially, blends of chemical stimulation of the tongue, palate, and nasal sense organs.
117. *Distancia*—Tendency to have feelings of distance with respect to surrounding objects or social stimuli.
118. *Temporal disorientation*—Tendency to have a poor sense of time. Exhibited by a false sense of time intervals, distorted temporal memories, and so on.
119. *Olfactorexis*—Tendency to obtain great satisfaction from many different types of odors and from situations in which they occur.
120. *Test objectivity*—A measure of how realistically the examinee has answered the test items. A low score usually indicates an automatic (not deliberate) tendency to make a favorable social impression.

121. *Auditory orexis for non-musical sounds*—Tendency to enjoy non-musical disturbances and to seek situations in which they occur.
122. *Food aversiveness*—Tendency to dislike foods, sometimes manifested as a disgusting or sickening dislike accompanied by a poor appetite.
123. *Sexorexis*—Tendency to be sexually appetitive, that is, to have a strong desire for sexual stimulation.
124. *Varietal love interests*—Feeling a need to love, not merely one, but several different persons.
125. *Misanthropy*—Tendency to feel disappointed or disgusted with the human race; in extreme form, manifested as hatred of the human race.
126. *Transsexual identification*—Tendency to identify with the cultural norms of the opposite sex. Transvestism is one outcome of this tendency.
127. *Masochism*—Tendency to experience sexual pleasure from being hurt. Unlike algolagnia, masochism is always sexual in nature (cf 57, 59, 73).
128. *Anti-intracception*—Tendency to resist attention to one's own feelings or thoughts, and to be against "prying" or concern about the feelings or thoughts of others.
129. *Primary suggestibility*—Tendency to be immediately induced to act in a specific way in response to the behavior of others.
130. *Sense of humor*—Tendency to enjoy the comic aspects of living situations, and be humorous in one's social relationships.
131. *Test consistency*—Tendency to respond in the same way to test items which measure the same personality trait.
132. *Test sincerity I*—Tendency to answer sincerely items in which *high* response loadings correlate positively with cultural approval. For example, the item "I finish my work before I play or rest" refers to culturally approved conduct. A response of *Always* would be a culturally approved response to this item, but may not be true.
133. *Test sincerity II*—Tendency to answer sincerely items which refer to culturally disapproved conduct. *Low*, instead of *high*, response loadings correlate positively with cultural approval. For instance, to the item "I feel that I have to set fires", a response of *Hardly ever* could be a culturally approved answer (cf 132).
134. *Self-depreciation*—An automatic (not deliberate) tendency to present an unfavorable view of oneself. Should be distinguished from a conscious, overt tendency to depreciate oneself (cf 55).
135. *Cholinergy (superfactor C)*—Refers to a factor common to a syndrome of primary traits. On the basis of theoretical considerations this superfactor has been called cholinergic, since high scores seem to indicate parasympathetic dominance, that is, a resilient tendency to feel energetic, happy, enthusiastic, secure or optimistic. For further details, see "The neutral theory of personality .....", *Trans. N. Y. Acad. Sci.*, 17, 315-30.
136. *Adrenergy (superfactor A)*—A factor common to a large syndrome of primary traits. Referred to as adrenergic, because high scores seem to indicate sympathetic dominance, that is, a tendency to feel sluggish, dejected, inferior, anxious, tense, shy, insecure, distractible, habitually fearful, or hypersensitive (cf 135).
137. *Self-acceptance*—Tendency to like yourself as you are, or more accurately, as you think you are. A measure of the acceptance of the real self.
138. *Adjustmental discrepancy*—The difference between the concept of the real self and the concept of what one would like to be. A measure of discrepancy between the real self and one's ideal self (ego-ideal).



139. *Authoritarianism*—Tendency to submit uncritically and devotedly to an idealized leader, but to aggressively seek out, dominate, condemn, or punish those who fail to conform rigidly with the ideals or dictates of the leader. An extreme radical authoritarian would be violently opposed to conventional, middle class values. An extreme conservative authoritarian would tend to be very hostile to those who fail to conform with conventional, middle class values.

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. BERGER, E. M. The relation between expressed acceptance of self and expressed acceptance of others. *J. Abnorm. & Soc. Psychol.*, 1952, **47**, 778-82.
3. BILLS, R. E., VANCE, E. L., & MCLEAN, O. S. An index of adjustment and values. *J. Consult. Psychol.*, 1951, **15**, 257-61.
4. BORGATTA, E. F., & ROBBIN, J. Some modifications of high speed computers for the methodology for social research. *Soc. Sci. Res.*, 1959, **43**, No. 4.
5. BURT, C. L. *The Factors of the Mind*. New York: MacMillan, 1940.
6. CATTELL, R. B. *The Description of Personality*. Yonkers: World Book, 1946.
7. ———. The description of personality: basic traits resolved into clusters. *J. Abnorm. & Soc. Psychol.*, 1943, **38**, 476-507.
8. ———. The description of personality: principles and findings in a factor analysis. *Amer. J. Psychol.*, 1945, **58**, 69-90.
9. ———. *Factor Analysis*. New York: Harper, 1952.
10. FAGIN, H. T. An investigation of the nature and magnitude of the components of retest unreliability. Doctoral thesis. New York Univ., 1950.
11. JENKINS, T. N. Some contributions to a neutral theory of personality. *Trans. N.Y. Acad. Sci.*, 1950, **13**, 9-12.
12. ———. The neutral theory of personality: new evidence and a revision of the argument. *Trans. N.Y. Acad. Sci.*, 1955, **17**, 315-30.
13. ———. How Well do you Know Yourself. *Global Form A42*, 1962, T. N. Jenkins, 21 Wash. Pl., N.Y. 3, N.Y.
14. ———. The second order components of human personality. *J. Psychol. Studies*, 1962, **13**, (in press).
15. KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
16. LENTZ, T. F. *The C-R Opinionnaire*. St. Louis: Character Research Foundation, 1935.
17. NEUHAUS, J. O., & WRIGLEY, C. The quartimax method: an analytical approach to orthogonal simple structure. *Brit. J. Statist. Psychol.*, 1954, **7**, 81-91.
18. RUNDQUIST, E. A., & SLETTTO, R. F. *Personality in the Depression*. Univ. Minn. Pr., 1936.
19. SMITH, P. A. A factor analytic study of the self-concept. *J. Consult. Psychol.*, 1960, **24**, 191.
20. SPEARMAN, C. *The abilities of man: their nature and measurement*. London: MacMillan, 1927.
21. THOMSON, G. H. *The Factorial Analysis of Human Ability*. Univ. London Press, 1950.
22. THURSTONE, L. L. *Multiple-factor analysis*. Univ. Chicago Press, 1947.
23. ———. The dimensions of temperament. *Psychometrika*, 1951, **16**, 11-20.

24. TRYON, R. C. Reliability and behavior domain validity: reformulation and historical critique. *Psychol. Bull.*, 1957, **54**, 229-49.
25. WHERRY, R. J., & GAYLORD, R. H. The concept of test and item reliability in relation to factor pattern. *Psychometrika*, 1943, **4**, 247-64.

*Department of Psychology*  
*New York University*  
*Washington Square*  
*New York 3, New York*



## DIFFERENTIAL IMPAIRMENT IN MULTIPLE SCLEROSIS\*

*Department of Psychiatry, Cornell University Medical College, and Hunter College*

---

CHARLES A. KNEHR<sup>1,2</sup>

---

### A. INTRODUCTION

Various aspects of cognitive impairment have been under investigation by this laboratory for several years. Interest has centered on clarifying the nature and patterns of losses to be observed in psychopathological illnesses, cerebral organic conditions and possible differentials between them.

In addition to studies of a small group of patients with Wilson's Disease (5) and a larger group of organics with assorted diagnoses (7) an opportunity was presented to examine a group of out-patients with the diagnosis of multiple or disseminated sclerosis.

While these patients fall into the same diagnostic category, it is known that the symptoms of multiple sclerosis vary extensively from one patient to another even though the pathological process is apparently distinctive. Superimposed is the episodic nature of the disease and wide variations in its time course. The gray matter tends to be spared while the sclerotic islets invade the white matter and are generally widely scattered from around the ventricular system to the cord. In some instances, the disease occurs as one of the features of a broader symptom complex. In its terminal stages, patients often exhibit signs of organic dementia.

At earlier stages of the illness, when patients may still be functioning fairly well in their daily activities, there are often motor and speech disturbances indicating defective motor control. These and other symptoms such as the frequently observed visual field defects, heightened tractability and mood characteristics ranging from depression to an overly cheery attitude suggest alterations of psychological functions before the disease process is well advanced from the standpoint of clinical evaluation.

The test battery in use in this laboratory seemed adequately suited to a

---

\* Received in the Editorial Office on June 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> From the Department of Psychiatry, Cornell University Medical College, The New York Hospital (Payne Whitney Psychiatric Clinic), and Hunter College.

<sup>2</sup> The interest and aid of Dr. Harold R. Wainerdi and his staff of the Demyelinating Disease Clinic of the Hospital For Special Surgery, New York, in making patients with multiple sclerosis and relevant clinical data available is most gratefully acknowledged.

search for the patterns of impairment in patients with multiple sclerosis of moderate severity. In respect to abilities, extension could be made beyond the more usually studied abstract intellectual functions since, in addition to these *g* functions, the tests of this battery are now known to load at moderate to high levels on at least five additional independent factors.

Thus equipped, the study sought to determine the severity and patterns of impairment in out-patients with multiple sclerosis and to compare such deficits as were found with the distribution and extent of losses in patients with clinical indications for diffuse cortical involvement.

## B. METHOD

### 1. Subjects

The anchor group for the study were patients ( $N = 11$ ) with a diagnosis of multiple or disseminated sclerosis, all but one of whom were being seen regularly as out-patients in the Demyelinating Disease Clinic of The Hospital for Special Surgery. All were symptomatic to some degree when tested and the mean duration of illness for the group was 8.5 years. From vocabulary performance and educational history (a mean of 14.6 years of schooling) it was estimated that all these patients were average or higher in preillness intellectual ability. Several were gainfully employed.

For comparison purposes, two other groups were drawn from the pool of control subjects and patients with diagnoses of cerebral organic conditions other than multiple sclerosis. These were selected on the basis of sex, age, and vocabulary score (Wide Range, Form B) without reference to other performance data. The mean ages for the controls (C), multiple sclerotics (MS), and organics (O) were 40.9, 41.2, and 41.5 respectively. The mean vocabulary scores were 83.8, 80.8, and 80.2 corresponding to vocabulary ages between 19 and 20 years. There were six men and five women in each group.

### 2. Tests

Out of a total battery of 10 tests, six were found in previous work to yield few false positives among normals while abnormal scores, when obtained, were generally found in the brain damaged (7). Performances on these six more successful tests were chosen as the basis for comparisons between the three subject groups of the present study. The tests<sup>3</sup> were:

*a. Progressive Matrices* (1938). This is a non-language test of the ability to discriminate complex patterns and to deduce relations and correlates. Scoring was in terms of the absolute number of problems solved correctly.

<sup>3</sup> Additional details on these and general procedure may be found in a previous article (6).



*b. Opposites association.* The easy to difficult stimulus words were given orally and timed manually with the scoring relative to both time and appropriateness of response.

*c. Flicker fusion test set.* The apparatus (after Landis) was designed to measure the critical fusion frequency (CFF) in cycles/second at the first appearance of a flickering field as the frequency declined from a level at which no flicker was reported.

*d. Series choice reaction time.* This device was arranged to provide a series of choice response situations, each initiated without delay by the preceding correct response of pressing the button corresponding to the light signal flashed on. Each sequence provided 20 trials in prearranged random order. The score was the mean of 60 trials.

*e. Pursuit rotor.* This apparatus was used to derive a learning curve for the accuracy of maintaining contact on a small rotating target with a hinged stylus. Twenty 15-second trials were alternated with 15-second rest periods. The score on each trial was the number of seconds during which contact was maintained.

*f. Divided attention.* The already practiced Pursuit task was performed alone. A second task, which required keeping a key depressed with the non-preferred hand when a 400 cycle tone of irregular duration and intervals was heard through a head set was also performed separately. The score was the change in performance when the two tasks were combined for an equal number of trials.

### C. RESULTS

It is evident in Table 1 that, on all tests except Divided Attention, the multiple sclerosis and organic group performances were inferior to the controls. It is equally apparent that, on tests which measure abstract conceptual activity (Matrices and Opposites), the performance of multiple sclerotic patients was closer to that of the control group than to the observed efficiency of the organics.

The implications of these data are more readily seen by reference to Table 2. When organic losses are compared to control data (C-O), all differences are significant. The multiple sclerosis group is significantly superior to the organics on the Matrices, occupies an intermediate position which is not significantly different from either controls or organics for Opposites while, like the organic group, they are clearly impaired on the remaining tests.

Taken together, these data indicate that for quite some time during the

TABLE 1  
A COMPARISON OF THE PERFORMANCE OF CONTROLS (C), MULTIPLE SCLEROTICS (MS), AND ORGANIC PATIENTS (O)  
ON THE INDICATED TESTS

Tests	Group means			Standard deviations		
	C	MS	O	C	MS	O
Progressive Matrices raw score	50.2	46.2	32.4	5.9	6.8	7.8
Opposites score	61.9	56.3	48.1	10.1	10.4	8.1
CFF in cycles/second	44.4	32.4	35.3	4.7	5.1	7.9
Series choice reaction time in seconds	0.77	1.23	1.03	.06	.25	.25
Pursuit learning:						
Mean of trials 1-5 in seconds	5.4	1.6	3.0	1.9	1.6	2.2
Mean of trials 16-20 in seconds	8.9	2.9	4.8	2.2	2.3	3.3
Divided attention loss in per cent	48.3	51.3	48.5	19.5	31.3	33.7



TABLE 2  
THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN GROUP MEANS  
FOR THE INDICATED GROUP PARINGS  
(A dash indicates that the significance did not reach the .05 level.  
Entries shown as .01 indicate this level of significance or better.)

Tests	C-MS	MS-O	C-O
Progressive Matrices	—	.01	.01
Opposites	—	—	.01
CFF	.01	—	.01
Series choice	.01	—	.01
Pursuit learning:			
Trials 1-5	.01	—	.01
Trials 16-20	.01	—	.01
Divided attention	—	—	—

course of the disease, patients with multiple sclerosis show preservation of abstract intellectual (*g*) ability. At the same time, activities which would be affected by limitations of visual efficiency and coordinated motor control tend to a slight but non-significant increase in impairment over that observed for organics.

Although the patient groups are small, the variation in individual losses on Matrices performance and duration of illness in multiple sclerotics is of interest. No relationship was observed in the present data.

The superior efficiency of the multiple sclerosis group over the organics occurred at all levels of difficulty on the Matrices. Figure 1 shows that the multiple sclerotics roughly approximate the performance of normals, while the organic patients are distinctly lowered in performance relative to these groups with losses most marked on test items of greatest difficulty.

Significant differences have been observed previously between control subjects and patients with a variety of cerebral histopathological or neurophysiological conditions on Divided Attention. Individually, over 60 per cent of such organics have shown marked losses. Generally, the larger loss occurs on the Pursuit task which has to be continuously monitored rather than on the Auditory task which, unfortunately, may be monitored discontinuously without excessive loss in moderately impaired patients. The initial pursuit performance by the MS group was generally so poor that there was little opportunity for an attentional deficit, if present, to be shown. The organics were also significantly inferior to controls on the initial Pursuit trials. But percentage-wise, the losses from separate to combined performances were equal in all three groups.

#### D. DISCUSSION

Forty years ago, Brown and Davis (2) asserted that the mental symptoms in multiple sclerosis are generally not like those associated with other organic

brain diseases. While wide variations were noted in type and severity of symptoms in multiple sclerotics, profound intellectual deterioration was said to be observed only occasionally, although some impairment seemed probable in a majority of cases.

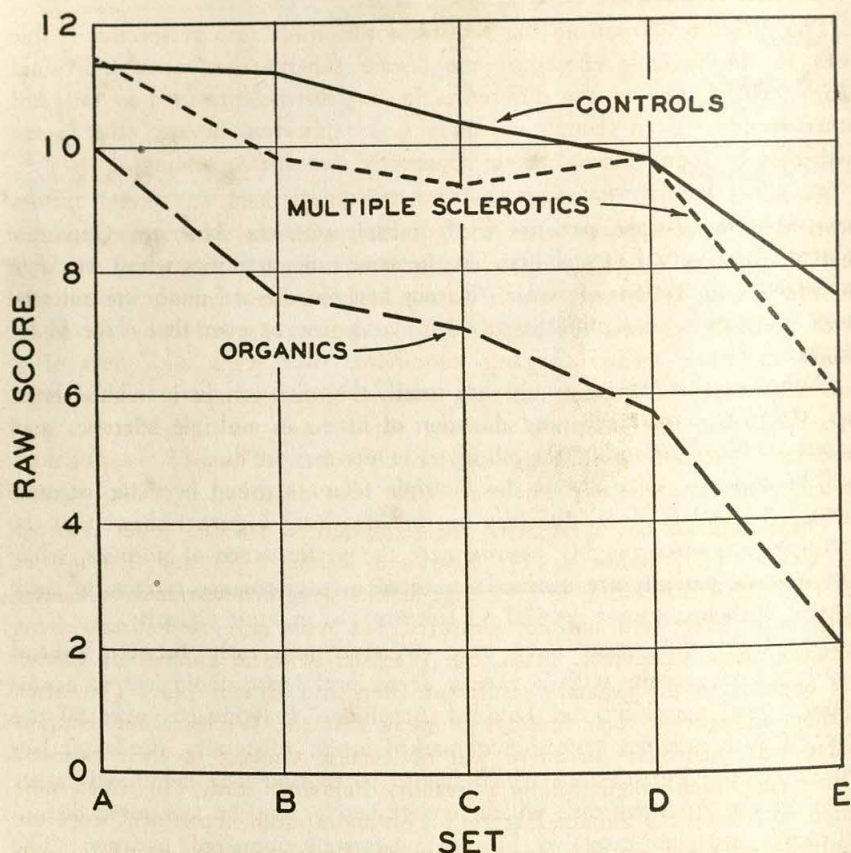


FIGURE 1

A COMPARISON OF THE MEAN SCORES OBTAINED BY THE CONTROLS, MULTIPLE SCLEROTICS AND ORGANICS ON THE FIVE SUCCESSIVE SETS OF TWELVE INCREASINGLY DIFFICULT PROGRESSIVE MATRICES PROBLEMS

In more recent years, there has been a growing number of investigations seeking to quantify the nature and extent of intellectual impairment in multiple sclerotics. The findings range all the way from the high proportion of significant losses on the Halstead Battery and Rorschach as reported by



Ross and Reitan (10) to the failure by Diers and Brown (4) to find significant changes when using the Wechsler-Bellevue impairment index. There are additional studies such as those by Baldwin (1) and Parsons, Stewart and Arenberg (9) which indicate slight to moderate losses in abstract intellectual efficiency.

The difficulty of putting the results of this study into perspective is due both to the variable effects of the disease superimposed upon individual characteristics, and to the differences in test instruments used in this and other studies. Even though no study is a replication of any other, some path may be found through these apparently inconsistent findings.

As noted before, the tests used for this study load on several factors other than what appears to be the *g* factor, with the Matrices, Opposites and Vocabulary (used for matching groups) showing substantial loadings on the latter. With tests which do not load on the *g* factor, the patients with multiple sclerosis are impaired to a degree approximating the losses found in other cerebral organic conditions. But it is with tests of *g* functions that the present multiple sclerosis group is clearly much better preserved than the organics, showing slight losses on the Matrices and moderate, but not significant, losses on Opposites. Thus, of the functions tested, general intellectual efficiency seems least disturbed in these patients.

These data do not agree with the significant abstraction losses reported by Canter (3), Baldwin (1), Ross and Reitan (10) and Parsons, *et al.* (9). It is interesting to note that deficits were demonstrated by Canter with both direct and indirect measures, but with only modest agreement between these techniques. Also, the Wechsler-Bellevue findings of Canter are opposed to the negative results of Diers and Brown (4). The latter authors interpreted their findings to mean that the Wechsler deterioration index was inadequate or there was no cortical damage in their patients. Other conclusions might follow as readily. Baldwin's study (1) which indicated no loss on the Shipley-Hartford Test and a mean impairment score for multiple sclerotics on the Hunt-Minnesota Test involved a control group with a mean T-score within one point of the borderline between control and impaired performances on this test suggesting a sizeable number of false positives.

The present findings of a sharp reduction in CFF from normal levels is in general agreement with the data of Parsons and Miller (8) and Titcombe and Willison (11). Since flicker thresholds are often significantly lowered in other cerebral organic diseases, and in view of previous work on central

vs. peripheral factors in CFF, there is no attempt here to evaluate the suggestions drawn from these studies as to the locus of the malfunctioning structures.

Taken together, the present data indicate a variety of detectable losses in patients with multiple sclerosis which are similar to those found in other organic conditions. At the same time, they tend to agree with the view put forth by Brown and Davis (2) that these patients differ considerably from other organics—at least with respect to the frequency of relatively well preserved ability to discriminate complex visual patterns, perceive new relations and solve abstract reasoning problems based upon these perceived relations. This difference is further attested by our findings of at least 80 per cent abnormal individual performances by a previous group of 50 organics and the present group of 11, while only two out of 11 multiple sclerotics made such scores on the Progressive Matrices. Whether such preservation of abstract intellectual ability will be an advantage to the practitioner who treats these patients is a question that may be answered only by relevant study.

#### E. SUMMARY

Three matched groups consisting of control subjects, ambulatory patients with multiple sclerosis, and patients with diagnoses of cerebral organic conditions other than multiple sclerosis were examined with a battery of tests previously found useful in detecting cerebral organic damage.

The multiple sclerotics were significantly different from patients with other neurological conditions in the frequency of rather well preserved abstract intellectual ability in the presence of losses in other tested functions that equalled or exceeded those of the organic group.

#### REFERENCES

1. BALDWIN, M. V. A clinico-experimental investigation into the psychological aspects of multiple sclerosis. *J. Nerv. Ment. Dis.*, 1952, **115**, 299-343.
2. BROWN, S., & DAVIS, T. K. Mental symptoms of multiple sclerosis. *Arch. Neurol. Psychiat.*, 1922, **7**, 629-634.
3. CANTER, A. H. Direct and indirect measures of psychological deficit in multiple sclerosis. *J. Gen. Psychol.*, 1951, **44**, 3-50.
4. DIERS, W. G., & BROWN, C. C. Psychometric patterns associated with multiple sclerosis. I. Wechsler-Bellevue patterns. *Arch. Neurol. Psychiat.*, 1950, **63**, 760-765.
5. KNEHR, C. A., & BEARN, A. G. Psychological impairment in Wilson's disease. *J. Nerv. Ment. Dis.*, 1956, **124**, 251-255.
6. KNEHR, C. A. Thinking impairment in acute schizophrenic illness. *J. of Psychol.*, 1962, **53**, 15-28.
7. ———. Psychological assessment of differential impairment in cerebral organic conditions and in schizophrenics. *J. of Psychol.*, 1962, **54**, 165-189.



8. PARSONS, O. A., & MILLER, P. N. Flicker fusion thresholds in multiple sclerosis. *Arch. Neurol. Psychiat.*, 1957, **77**, 134-139.
9. PARSONS, O. A., STEWART, K. D., & ARENBERG, D. Impairment of abstracting ability in multiple sclerosis. *J. Nerv. Ment. Dis.*, 1957, **125**, 221-225.
10. ROSS, A. T., & REITAN, R. M. Intellectual and affective functions in multiple sclerosis: A quantitative study. *Arch. Neurol. Psychiat.*, 1955, **73**, 665-677.
11. TITCOMBE, A. F., & WILLISON, R. G. Flicker fusion in multiple sclerosis. *J. Neur. Neurosurg. Psychiat.*, 1961, **24**, 260-265.

*Payne Whitney Psychiatric Clinic*  
*The New York Hospital*  
*525 East 68th Street*  
*New York 21, New York*

## A DISCUSSION OF CHRONOGEOMETRY AS RELATED TO NON-STATIONARY TIME SERIES SUCH AS THE PHYSIOLOGICAL ELECTROGRAM\*<sup>1</sup>

*Columbus Psychiatric Institute and Hospital*

---

SAMUEL J. M. ENGLAND AND BENJAMIN PASAMANICK

---

The essence of the problem in psychological physiology is the observation of changes in measurements which are notably variable. The analysis of such data may require techniques that take cognizance of the fact that the usual notions as to the meaning of *time* can be misleading when applied to the biological system; events do not occur uniformly throughout the various parts of the animal's nervous and physiological activity. This inherent condition imposes upon the experimenter the necessity to resort to special techniques for the analysis and understanding of multiple interacting non-stationary time series.

The following proposal was briefly discussed by England and Pasamanick (5); not every detail of the proposal is reproduced here. It might be of interest to outline some of the circumstances preceding this proposal. In November 1960 we came upon a report of Scientific Congresses and Conferences by Menitskii (6) in translation. The following excerpt may be found on p. 129 of the translated edition:

Much attention was given at the Conference to questions connected with cardiovascular investigation and diagnosis. I. G. Akulinichev and E. B. Babskii, in conjunction with the engineers of the Computer Institute, State Planning Commission of the U.S.S.R., presented two papers. The first described a new apparatus, an analyser of the electrical activity of the heart. In this apparatus the electrocardiogram is divided into separate cycles, the limits of which are the peaks of the waves. These cycles are recorded one on top of the other, which facilitates analysis of the electrocardiogram and makes for economy in prolonged recordings. The electrocardiogram obtained in this manner is reproduced on the screen of an electron beam apparatus with prolonged afterglow, and is simultaneously recorded on the photographic strip. On the evidence

---

\* Received in the Editorial Office on June 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The present essay is a slight amplification of a version read before The 4th International Conference on Medical Electronics, July, 1961, sponsored by the Institute of Radio Engineers, New York 21, New York.



of the authors, this method is most useful for investigation of fluctuations of the cardiac rhythm under normal or pathological conditions.

Contemplation of this method of Akulinichev and Babskii resulted in our inquiry as to the suitability of a slight further innovation of the kind reported here. Details leading to the present formulation cannot be included here.

For our proposed method of recording and monitoring the two physiological variables, Figure 1 conveys the impression to be discussed. The two variables of interest are monitored electrically and recorded in epochs of predetermined duration. The lines of the record may display both physiological entities together, or the *accented* or reference variable may be suppressed, as is the case in our illustration. The rate of trace for each line or duration can be continually calculated by the apparatus, interval by interval as the record evolves. The experimenter may observe transitions in these rhythms almost mechanically and apprehend immediately the physiological processes with no more than a minimum of explicit analysis, assuming that the experimenter can provide the necessary background of theoretical abstraction as the basis for interpretation. It would perhaps not be so important that the actual hardware be constructed for effecting this form of display except that this form of demonstration may produce such "visual compulsion" which can carry over to the *logical* compulsions of geometry, even though one may then choose to renounce the visual display in favor of more complex analytical procedures. Figure 2 exhibits the geometrical model of our discussion.

It may be profitable to undertake an examination of some of the premises underlying this method of conceiving physiological processes. Space and time is said to have its origin in the relations between events. Physicists have usually conceived the physical world and the *ether* of the physical world in a manner very like Newton's conception, but the more recent formulations of physics have notably replaced the material ether or *space* by a complex of events and their characteristic activity with respect to one another: as it were, an "ether of events." For our purposes here we are asked to conceive of *space* as the exhibition of internal relations between physiological *time-systems*. The *system* is the unit of significance in our analysis. As an accretion to our discussion there are certain analogies and homologues to the visual conceptions of space which may be of more than mnemonic significance.

Figure 1 will reveal something of the practical considerations that are encountered in this form of visual display: it can be noted that as the situation proceeds from the bottom upward, the record below the letter B is adequately stationary; in this case there are eleven heartbeats underlying the record of

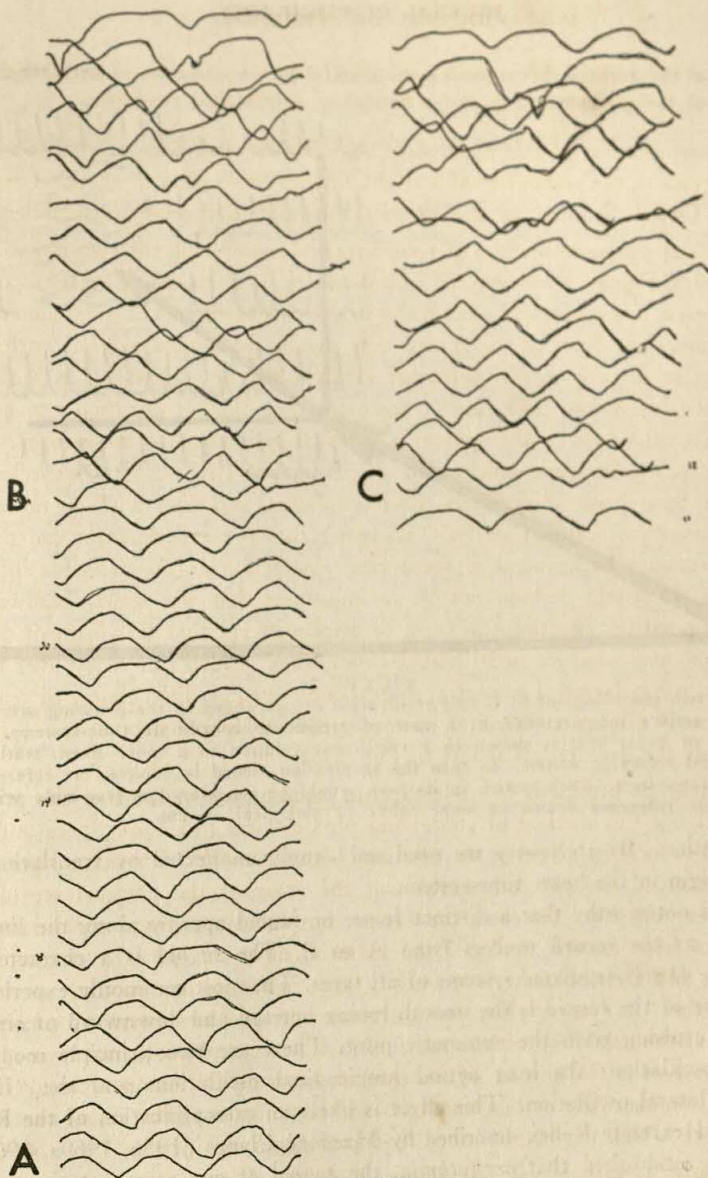


FIGURE 1

The successive lines of the record above are tracings from a conventional pen recording device. The electrical analog of respiration was recorded by means of a transducer strapped around the subject's chest. The upward and downward excursions of the traces represent inhalation and exhalation movement, connected with the subject's breathing. The record proceeds historically from the bottom toward the top. The successive lines were determined as to length by the lapse of 11 beats of the subjects heart recorded simultaneously in the usual manner for ECG.



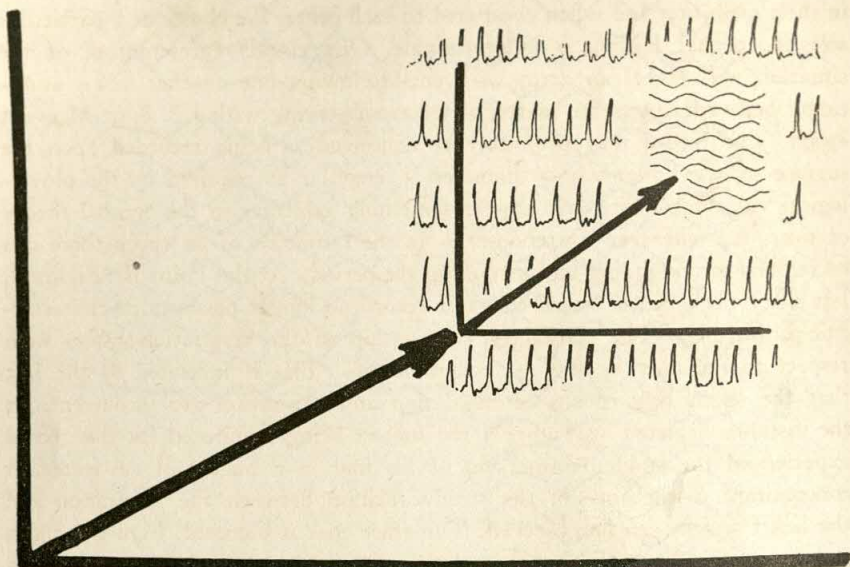


FIGURE 2

The two recordings of ECG and respiration are arranged in the drawing according to an artist's interpretation of a nest of rectilinear coordinate time-systems. The record of heart beat is shown in a rectilinear relation to a more fixed, stationary or closed reference system. In turn the respiration record is depicted as referred to the heart-system. Each system in its turn would be considered a free-body with its principle reference frame in some other physiological system.

respiration. By *stationary* we need only imply unaffected by translations of the origin in the heart time-system.

It is noteworthy that a distinct fugue or fugoid appears along the lines of zeroes as the record evolves from A to B. The fugoid is a characteristic feature of self regulated systems of all types. The most commonly experienced instance of the fugoid is the smooth loping upward and downward of aircraft when cruising with the automatic pilot. There are two principle modes of such oscillation: the long period longitudinal oscillations and the "Dutch Roll" lateral oscillation. This effect is likely an exemplification of the Respiration-Heart-rate Reflex described by Manfred Clynes (1956, 1960a, 1960b). Clynes established that *arrhythmia*, the fugoid of our representation of the situation, is initiated by *stretch receptors* within the thorax, rather than by hemodynamic or central factors.

It is fundamentally important that insofar as our record reveals the two

time-systems to be stationary, both when compared to themselves at intervals in their evolution and when compared to each other, the choice of a particular size of epochal folding is indeterminate. Our visual representation of the situation will reveal an array of events following one another in an undistinguished orderliness; no matter if we are observing with a 2, 3, or M event epoch. The record may profitably be conceived as being recorded upon the surface of a cylinder whose diameter is irregular as required by the physiological variable. There may appear something arbitrary in the epochal theory of time, for wherever you choose to fix the terminals of an epoch there can be reasons for extending or contracting the period. At the Point B (Figure 1, left side) there can be noted that our record no longer possesses its characteristic uniformity. The stationary relationship of the respiration-system with respect to the heart-system no longer obtains. This is indicated by the fact that the succeeding epochs of inhalation and expiration are incoherent. In the instance depicted by Figure 1 the subject being monitored for this record experienced the sudden immersion of his foot in a bucket of ice water. A concomitant disturbance in the steady relation between the respiration and the heart system can be observed. The trace that is pictured, Figure 1 right, starting at C, illustrates that the relation we wish to observe can be recovered by the expedient of shifting to a basis of 12 heart beats in the heart time-system. However, the *line-of-zeroes* of the respiration system remain skewed at a sharp angle to the right in the record, and it would appear that 13 or 14 units of the reference system would be necessary to recover the stationary situation. Figure 3 illustrates that the total number of epochal units is somewhat arbitrary for viewing; but the fundamental unit of epoch length is not arbitrary. The record depicted on the left, Figure 3, is the same as that depicted in Figure 1, left. Now, however, the basis for each epochal line is 22 units in the heart time-system. The fugoid is again evident and the contingency of ice water immersion can be noted to affect the upper part of the record. Figure 3, right, illustrates that 24 units in the heart system will not secure a visually stationary pattern with which to discriminate the behavior of the respiration-system with respect to the heart-system.

Another aspect of our topic bears examination however. There is the hint of a condition that deserves investigation in the left record of Figure 3, i.e., that the basis of eleven for generating the record of this particular subject, in this particular condition, might be shown to *remain* an effective basis for recording if the record were of sufficient epochal width (say 44 or 88 units) as well as, of course, length. It may prove that the perturbing effect of submerging the foot in ice water would appear as an anachronism visibly con-



tained or circumscribed on all sides by a uniform record. This would serve as an example of the possibility of a *focal region* within the larger group of time independent recurrences. There are certain geometric considerations underlying this reasoning that leads us to contemplate the physiological electrogram in this way.

In principle any periodic process can serve as a base for a time measurement. Some regularity of recurrence is essential, otherwise measurement would be impossible. Beats of the heart recur, breathing recurs. To describe any phenomenon requires a definite frame of reference or base, either explicit or implied. A reference or base is significant only when compared with a change of something that is present along with the reference frame. Something that has the character of *change* can be described only with reference to the stationary or referent elements present along with the changes. It may be relevant in this connection to recall the remarks of Einstein (4) concerning time and *stationarity*: "It is not possible to obtain a reasonable definition of time with the aid of clocks which are arranged at rest with respect to the body of reference."

It is not altogether irrelevant, although possibly inappropriate, at this particular connection to recall that members of the same family distinguish one another by their *differences*; and cannot as a rule discern the similarities in their appearances. In the usual context persons not themselves members of that family may be more impressed with the *similarities* among the members, and perhaps only secondarily impressed with the differences. This would seem compatible with the notion that the perception of other *time-systems* is facilitated through *transitions* with respect to similarities present in the situation along with the changes that are occurring. Changes that are immediately distinguished, are discriminated by comparison with time-systems that are going on in a stationary manner, and which serve as the connections of time-systems among themselves. When frames of reference or time-systems are considered together one of them may be more *inertial* with respect to others.

The geometry of a frame of reference or base is connected with events being measured in the reference space by a system of rules. At different points along the base we can measure the position of the objects in the space referred to the base. Figure 3 illustrates that our reference frame, the heart-time-system, is not a rigid scaffolding. This is evident on the record from the fact that the epoch lines of explicitly displayed inhalation and exhalation are of varying total length, although each epoch consists of 22 heart-beats. That is to say, if copies of two separated epochal lines of our record were

transported and compared locally, they would be different in length. Although it is permissible to envisage there being a clock at each of its epochs, all these clocks are not synchronous: some are faster, some are slower; time is a variable magnitude.

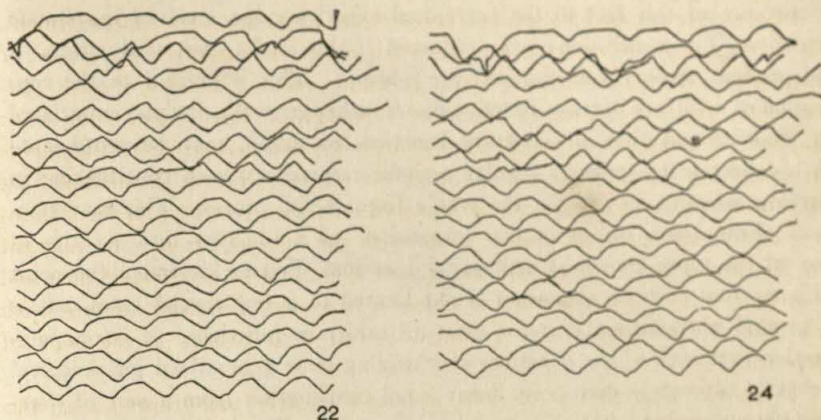


FIGURE 3

The respiration record of Figure 1 is shown here as recorded in epochs of twice the number of heart-beats as that of Figure 1. The record attempts to illustrate that the choice of epoch length is not completely arbitrary; there are discrete periodic lengths of the reference-system that result in the desired exhibition.

It is for this reason, and for other reasons that the notion of clocks is less useful to psychological physiology than the notion of *time-systems*. Among the other reasons is the habit of accepting the hands of the clock as the clock. The hands are largely irrelevant to the nature of a clock. Only when there is relative motion is there *time*. Perhaps a major significance of the relative theory of events is that mathematical operations which have been legitimized for space by rational analysis and found to be empirically useful in any number of instances may now also apply to *time*. When we speak of a clock we would mean anything which goes through some regular periodic performance or process, and each periodic process gives a measure of its own "proper" time. The "proper" time for an object is said to be time-order of events in its immediate neighborhood. Space is to be considered timeless only in the limiting case; where the relation between the variables (in this case, heartbeat and respiration) are fixed and unchanging. For there to be a timeless space the set of neuronal *event particles* constituting respiration must remain fixed to the assigned coordinates, while the epochal lines of the



reference system may have every possible integral value. That is to say, one *time-system* remains *stationary*, or time-independent with respect to the other.

We may now attempt to convey the significance and interrelations characteristic of time-systems as specifically derived for physiology. There is some agreement in the literature that nerve action is coded logarithmically. The importance of this fact to the perceptual experience has received inadequate attention. The conclusion seems enforced if this is the case, that effects of independent sensory contributions are additive. This is helpful to the conception of arbitrary nervous functioning. A rather general principle, enunciated by Fourier was that an arbitrary function, or graph, may be represented by a collection of series of regular periodic recurrences, each contributing in varying amounts to sum up the whole linear serial process. This mathematical phenomenon occurs in the "expansion" of a function into a series of any of the types known as orthogonal functions. Fourier defended the point of view that such an expansion is not limited to a few special cases, but is a general phenomenon true for most all arbitrary functions. It is our suggestion that each of the durations synthesizing some generalized physiological behavior ultimately derives its linear serial contribution from a *nest* of *time-systems* somewhat after the manner suggested by Figure 4. This family of time-systems is generative of a family of *durations* that have the two termed relation of *extending one over* the other; the first harmonic of a fourier expansion may be said to *extend-over* the second, and the second over the third, etc. Figure 5 may serve to convey the notion of a system C *extending over* systems H and R. The conception of an umbrella would misconstrue the proper of central significance of *extending over*; it would militate less our understanding to visualize a string that forms the axis upon which are strung a series of beads. There are disadvantages in such an image however, for the essence of the situation is the motion: *motion* is the thing. . . . It is the hope that by contemplating the situation of the physiological electrogram in this way, we may be led to adopt geometrical analogies whereby a large class of behavioral operations can be linearly related. Neural events of necessity would be conceived as a multidimensional vibrating system.

We would have the reader further consider some of the principles of functioning of the closed system of Figure 5. For certain limited purposes H, C and R could be thought of as geared wheels, whereupon any velocity of motion whatever would be possible for the complete system H, C, R with respect to the implied larger system including the observer external to system H, C, R. However, the system considered in itself, without additions, would within itself be perfectly rigid except on the assumption that some of the "teeth" in

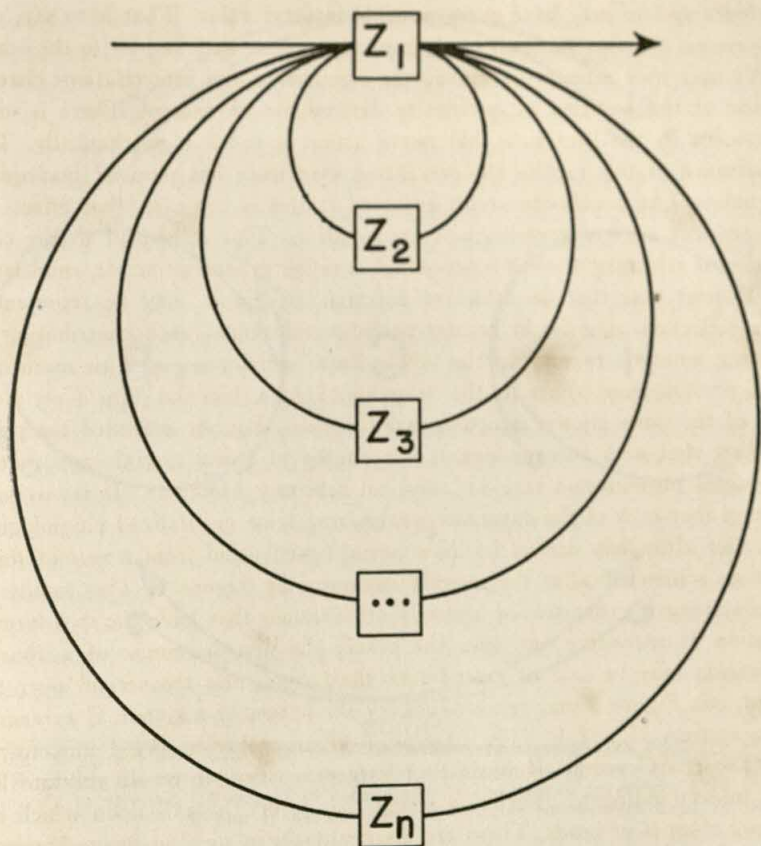


FIGURE 4

This symbolic arrangement of closed systems of activity would serve to emphasize that for any given nexus of input-output relationship there are a number of systems implicated, and these would have the characteristic of successively longer and longer intervals to complete one cycle, and might slip their relative motion at the nexus as if a succession of cylinders, end view.

the gear wheels were to slip. This slippage is the very essence of relative motion between time systems. Let us consider the matter from a slightly different viewpoint. Time-system C may be thought of as a transformation matrix or connection matrix between the two manifolds of activity  $\mathcal{H}$  and  $\mathcal{R}$ . Furthermore, given  $\mathcal{H}$  stationary with respect to  $\mathcal{R}$ , a transformation or transfer from time-system  $\mathcal{H}$  to time-system  $\mathcal{R}$  by way of  $\mathcal{C}$  would imply that  $\mathcal{C}$  is a group of linear constants. What is perhaps more important to nervous



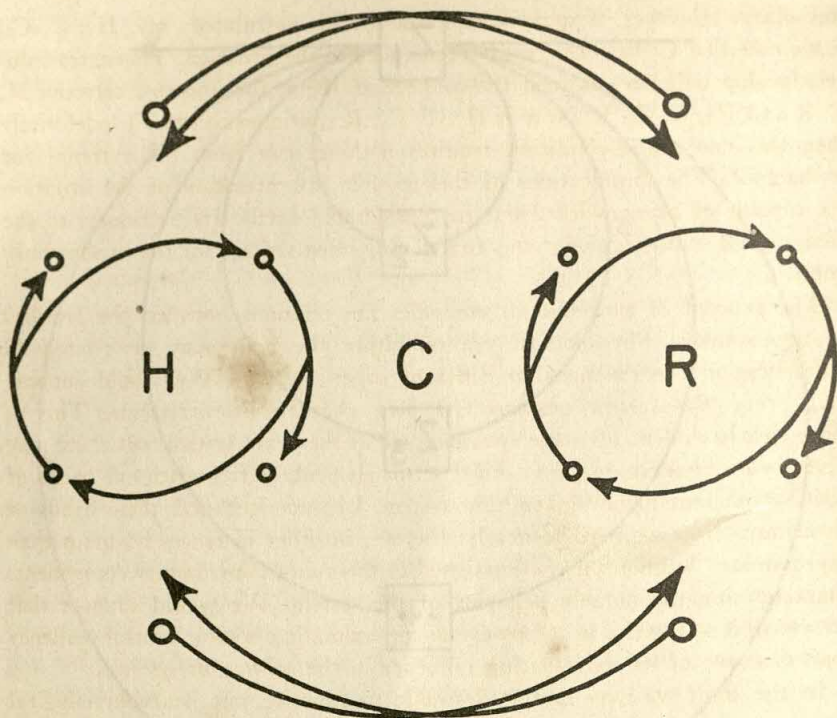


FIGURE 5

The symbolic arrangement of connections depicted above would suggest that a closed system of activity H would be related to a closed system of activity R by means of a system of connections C, extraneous to H and R. Activity of C would reflect or be a measure of relative motion between H and R.

action is that if C is by way of being some part of the brain, H is the matrix of nervous heart-system connections and R is the matrix of respiration-system connections, then C extends over them and perturbations of either H or R is reflected in C; and all of them together are not closed to systems external to H, C and R. When there is relative movement other time-systems extraneous to H, C and R are necessary to provide linear serial contributions between, as it were, H and C, and C and R. This influx takes place in proportion as there is relative motion between time-systems. In the absence of relative motion between time-systems the contributions from the extrinsic systems are latent.

By way of example, given some dynamic condition H, C, R there will be no necessity of any "outside contribution"; H, C, R will be complete and

stationary. However, if some second condition is postulated; say,  $H_1$ ,  $C$ ,  $C_1$ ,  $R$  we note that  $C_1$  has of necessity been added to the situation. This extraneous relationship will have formed the connection, if not the impetus, between  $H$ ,  $C$ ,  $R$  and  $H_1$ ,  $C$ ,  $C_1$ ,  $R$ . Now if  $H_1$ ,  $C$ ,  $C_1$ ,  $R$  continues to prevail indefinitely then this new static condition requires nothing new from the external for its balance. The ramifications of this possible interpretation of the situation are capable of being worked out in considerable detail. Its usefulness to the analysis and synthesis of nervous functioning must remain for the present only implicit.

The concept of time-systems embodies the relations between the families of time-systems. Physiological effects within the body may be related to differences in the reactions of different organ systems. We would suppose that every *time-system* possesses its own characteristic tropisms. This is more clearly evident for some systems, such as the heart system, but there may nevertheless be no more than a finite set of elementary frequencies or bands of activity inherent in any given time-system. In some instances these bands of elementary tropisms may be nearly discrete; in other instances no more than approximate "location" may be assigned to the major contributory components characterizing the notable behavior of the system. We would suppose that no system is so fugitive as to possess no approximating note or central tendency representative of its coordinating influence in the scheme of systems.

In the limit we may find it desirable to conceive our neurophysiological system as *co-ordinate species* of time-systems each limited to a narrow temporal aspect of our physiology. This is a simplicity not realized in actual practice, but this fact may not affect the usefulness of the concept as characterizing the ideal case of nervous functioning. There is general agreement in the literature related to classical or Pavlovian condition that the arrangement or distribution of nervous action within the cerebral cortex tends to form a fixed pattern which is sustained more easily in proportion to its repetition under the same circumstances. There comes into existence a dynamic stereotype or systematic pattern which can be maintained with less and less expenditure of nervous energy. This stereotype or fixed history develops an *inertia* which resists any alteration by new stimuli, or what we can call new *boundary conditions*. This condition of inertia or rigidity is sufficient to cause us to undertake an examination of the theory of objects, or rigid bodies.

Reichenbach (7) reminds us that physical objects change their shape and size only very little when affected by outside forces but that this relative stability is no ground on which to give a preference to solid bodies for the



definition of congruence. He suggests that the concepts *rigid* and *solid* should be distinguished one from the other. Rigidity may be said to characterize an object, where change of shape is small if the exterior forces are small relative to the interior forces. This definition involves the supposition of a closed system. While a completely closed system can never be realized, systems are nevertheless closed to a greater or lesser degree of completeness. A closed system contains interior forces and infinitesimal closed systems are rigid in the proportion that exterior forces of larger systems extending over them have more nearly vanished than interior ones. "This definition of the rigid body is not explicitly given in the literature of physics, but it is that definition on which the whole system of physics is based" (7, p. 23). It is our suggestion that the coordinative definitions for the purposes of psychological physiology may be related to the definition of congruence by means of more or less *rigid time-systems*. A more or less *rigid time-system* considered in this way has some of the properties of an object. Among the properties that specify an object is that it have among its relations that of *time-independence* over the interval. An object can never be specified without there being also specified a relevant future and a relevant past—for *that* object. The relevant future and a relevant past is usually implicit in the specification of an object. A more detailed examination of the theory of objects can be found in Whitehead (8, 9).

The method of observing physiological activities proposed here has the practical advantage of suggesting its relation to a body of reasoning having the highest abstractive significance. When further elaborated we may have a contribution to a system of theory in connection with *time*. Failing this the proposal may provide an impetus for the construction by others of a system of knowledge about time. One should not suppose that the rather general relative theory of time that was first developed in connection with electromagnetism and light, has only this restricted application. For the most part the subject has a rather secondary status in the field of physics. There is good reason for this, involving the fact that the particular circumstances of observation of physical events do not typically demand what is sometimes disparagingly called "relativistic correction." The observations are all made with respect to the same time-system, and "the comparative absence of applications leads to a general neglect."

#### REFERENCES

1. CLYNES, M. E. Simple analytical method for linear feedback system dynamics. *Trans. AIEE*, 1955, 55, 2, 377-383.

2. ———. Respiratory control of heart rate: Laws derived from analog computer simulation. *IRE Trans. Med. Electronics*, 1960, ME-7, 2-14.
3. ———. Computer analysis of reflex control and organization: Respiratory sinus arrhythmia. *Science*, 1960, **131**, 3396, 300-302.
4. EINSTEIN, A. "Relativity." London: 1920.
5. ENGLAND, S. J. M., & PASAMANICK, G. The geochronograph method for representing non-stationary electrograms, such as heartbeat and respiration. *J. Psychol.*, 1961, **52**, 149-155.
6. MENITSKII, D. N. Radioelectronics in physiological research (*Fiziol. zh. SSSR* **45**, No. 6, 740-745, 1959), *Sechenov physiol. J. USSR*, **45**, No. 6, 1959 (Pergamon Press).
7. REICHENBACH, H. "The Philosophy of Space and Time." Dover, S443, 1958.
8. WHITEHEAD, A. N. "An Inquiry Concerning the Principles of Natural Knowledge." Cambridge: 1919, Cambridge: 1955.
9. ———. "The Concept of Nature." Cambridge; 1920, Ann Arbor; 2A, 1957.

Research Division  
Columbus Psychiatric Institute  
473 W. 12th Avenue  
O.S.U. Health Center  
Columbus 10, Ohio



## FACTOR ANALYSES OF SCHIZOPHRENIC AND ORGANIC TEST DATA\*

*Department of Psychiatry, Cornell University Medical College, and Hunter College*

---

CHARLES A. KNEHR<sup>1</sup>

---

The data of previously reported impairment studies (2, 3) have received additional statistical treatment in the form of a principal components factor analysis and Varimax rotation of the resulting factor axes for each of the three experimental groups studied.<sup>2</sup>

These analyses were made to provide a check on the a priori assumptions for these investigations. These additional data do not alter the inferences already drawn, but lend support to the hypothesis relative to the differential nature of impairment in cerebral dysfunction and in acute schizophrenics when observed in the latter. Although two or three measures for each of five postulated factors were included in the test battery, the primary purpose of the study was not to explore the whole domain of impaired ability or to identify the psychological variables. Any such yield would be considered a bonus beyond a better descriptive analysis of the experimental findings.

Eight of the obtained principal factors were rotated although only six of these had latent roots in excess of 1.0 for schizophrenics and organics while there were seven for controls. The resulting factors for the three groups were matched both by inspection and by the use of a coefficient of congruence (1, p. 257). Significant matches among all three groups were obtained for six factors, resulting in a fairly high degree of factor pattern similarity between controls, schizophrenics, and organics. The remaining two rotated factors showed dissimilar loading patterns from one group to another and may represent error factors.

For present purposes, the factor patterns for schizophrenics and organics are presented in Table 1. The data for controls are omitted to simplify the discussion. These patterns show the loadings of nine out of ten tests on the six factors which could be matched between the groups. The data for the

---

\* Received in the Editorial Office on July 3, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> From the Department of Psychiatry, Cornell University Medical College, The New York Hospital (Payne Whitney Psychiatric Clinic), and Hunter College.

<sup>2</sup> The factor analyses and rotations were computed by the Educational Research Corporation, Cambridge, Massachusetts.

TABLE 1  
A COMPARISON OF THE FACTOR PATTERNS FOR ACUTE SCHIZOPHRENICS AND ORGANICS FOR THE INDICATED TESTS  
(Moderate to high factor loadings as shown in italics)

Test	Schizophrenics (Rotated factors)						Organics (Rotated factors)					
	<i>F</i> <sub>1</sub>	<i>F</i> <sub>2</sub>	<i>F</i> <sub>3</sub>	<i>F</i> <sub>4</sub>	<i>F</i> <sub>5</sub>	<i>F</i> <sub>6</sub>	<i>F</i> <sub>1</sub>	<i>F</i> <sub>2</sub>	<i>F</i> <sub>3</sub>	<i>F</i> <sub>4</sub>	<i>F</i> <sub>5</sub>	<i>F</i> <sub>6</sub>
Progressive Matrices	<i>.87</i>	—15	—06	26	01	00	<i>.71</i>	—11	—24	14	08	—24
Vocabulary	<i>.82</i>	23	01	18	06	—21	<i>.94</i>	10	—08	08	00	01
Opposites	<i>.59</i>	43	—23	33	—12	—31	<i>.88</i>	01	—02	—07	01	—08
Critical flicker frequency	06	<i>.91</i>	09	—01	06	08	05	<i>.88</i>	—14	14	03	14
Series choice reaction time	—32	—24	<i>.60</i>	21	—38	—03	—19	06	<i>.54</i>	03	—18	01
Loss—divided attention—pursuit	26	02	—07	<i>.90</i>	02	—16	04	01	—02	<i>.99</i>	—06	00
Loss—divided attention—auditory	00	03	01	25	<i>.86</i>	09	03	02	—04	07	<i>.98</i>	08
Span of apprehension	—12	14	—08	23	—07	<i>.79</i>	—10	—09	03	01	16	<i>.86</i>
Pursuit learning	06	09	—22	25	<i>.50</i>	08	10	10	—08	—10	—05	08
Persistence	—01	—17	— <i>.84</i>	25	—05	10	09	17	— <i>.89</i>	—05	—01	—03



test not shown (After-Image) were incomplete for all groups, especially the organics for whom scores were obtained on only 32 per cent of the patients.

The initial assumption that the Progressive Matrices, Vocabulary and Opposites would load primarily on the same factor is borne out. It seems reasonable to refer to these as primarily tests of  $g$  functions, especially in view of previous work with these tests. It seems that this is the only factor in these studies which can be given some identity of psychological meaning.

CFF, Series Choice (and After-Image) were chosen as possible tests of inertia of mental functioning or perseveration. Previous work here and elsewhere suggested that there may be no unitary factor of perseveration. If there is, the chosen tests do not indicate it since CFF shows a high loading on the second factor and on no other for either group while Series Choice is largely independent of this factor. The Series Choice as devised shows a moderate positive loading on the third factor on which Persistence shows a fairly high negative loading. Whatever these factors mean psychologically, it is clear that CFF and Series Choice are not measures of the same variable with the latter test undoubtedly having a complexity greater than one. From the present data there is no way of deciding what these tests are measuring, whether one or the other, both or neither are measuring an inertia component in their respective systems.

The loss on Divided Attention is made up of the combined loss on the two separate tasks. It is interesting to note that the Pursuit loss (Factor 4) was independent of loss on the Auditory task (Factor 5). Data not shown indicate the loss for the tasks combined correlates highly with Pursuit loss which contributed the major share to figure of total loss on Divided Attention.

In turn, the Span of Apprehension loads chiefly on the sixth factor, being largely independent of other variables. Pursuit learning was a measure of improvement and showed no consistent loading patterns in the three experimental groups.

Two measures of variability in amplitude of performance (Spearman's Oscillation) were included in the present analysis. Their patterns of loadings were dissimilar and occurred chiefly on factors contributing but little to the total variance.

The relative independence of most of these tests from each other and the tests of  $g$  function provide support for previous findings (3). It was suggested that organic factors provide lower limits on the functional possibilities inherent in a nervous system and would be likely to become manifest both in complex psychological activity as well as in simpler processes requiring less complexity of thinking activity. Further speculation held that losses in acute

schizophrenics would be in those activities involving complex central integrative activity subject to momentary variations in perceptual sets, perceptual conflict and varying motivational trends and not to abilities which would suffer primarily from a reduced capacity for performance.

While group results showed the controls and schizophrenics to be different from organics at equivalent levels of significance, individually the schizophrenics relative to the controls showed the highest percentage of loss on Opposites and the Progressive Matrices. The previous inference to the effect that our acute schizophrenics when impaired would show the effect chiefly in *g* functions achieves confirmation on the nature of the functions most impaired. It is now apparent that the tests on which the schizophrenics, like the controls, show infrequent abnormal scores are largely independent of the *g* factor, but what these tests are measuring is not known. In any case, the suggestion that non-*g* tests are the measures of choice in detecting organic damage among psychopathological features seems to be sharpened by present data.

In another study on a small group of multiple sclerotics (4) it was in tests of *g* function that these patients showed the greatest sparing. Here again the crucial measures in the detection of organic involvement in the individual case seem to be the tests independent of *g*.

Differentials in data such as these raise the question as to whether impaired performances represent quantitative or qualitative changes. In noting the similarity of the schizophrenic and organic factor patterns above, one is tempted, in the absence of indications to the contrary, to state that the organization of the abilities measured are qualitatively similar in both patient groups and that the observed variations are quantitative. This is perhaps a safe statement since the history of scientific progress shows that the qualitative regularly becomes quantitative whenever a system of numerical ordering can be applied to observations. It is still probably worth saying that acute schizophrenics differ radically from organics quantitatively along some dimensions and to a lesser degree along at least one other.

#### SUMMARY

Principal components factor analyses with Varimax rotations were completed on data collected previously in the study of impairment in schizophrenics and organics.

The factor patterns for the loadings of nine tests on six factors were similar for the separate experimental groups.

Some of the implications of these data for the initial assumptions and reported results were discussed.



## REFERENCES

1. HARMAN, H. H. *Modern Factor Analysis*. Chicago: Univ. Chicago Press, 1960.
2. KNEHR, C. A. Thinking impairment in acute schizophrenic illness. *J. of Psychol.*, 1962, **53**, 15-28.
3. ———. Psychological assessment of differential impairment in cerebral organic conditions and in schizophrenics. *J. of Psychol.*, 1962, **54**, 165-189.
4. ———. Differential impairment in multiple sclerosis. *J. of Psychol.*, 1962, **54**, 443-451.

*Payne Whitney Psychiatric Clinic*  
*The New York Hospital*  
*525 East 68th Street*  
*New York 21, New York*

## THE CONTINUITY-DISCONTINUITY CONTROVERSY IN PAIRED-ASSOCIATE LEARNING\*

*Department of Psychology, Clark University*

---

NELSON BUTTERS<sup>1</sup> AND MORTON WIENER

---

### A. INTRODUCTION

In a recent series of studies Rock and his associates (6, 7) have re-examined the continuity-discontinuity issue in the area of human learning, and presented evidence from which they inferred that paired-associates are learned in an all-or-none fashion and that repetition plays no role in the learning of the associations. In one paradigm, Rock (6) found that the learning of the list of paired-associates was not retarded, as would be expected if weak associations had been formed by replacing the missed items with new items in each succeeding test trial.

In another paradigm, Rock and Heimer (7) found that paired-associate items that had been presented three times and missed in two test trials were learned no more quickly than items that had never been presented previously. From these results, it was concluded that no weak associations had been formed during the three presentations, and therefore learning is all-or-none.

In other recent studies Estes (3); Estes, Hopkins, and Crothers (4); and Clark, Lansford, and Dallenbach (2) using somewhat different paradigms have essentially confirmed Rock's original findings.

Despite the dissimilarity of experimental paradigms, all of these cited studies infer discontinuity from a failure to find significant differences between their groups. That is, they all infer discontinuity from evidence which is based on the acceptance of the null hypothesis. It would not seem possible therefore to determine whether the absence of differences indicates no real difference or can be attributed to absence of adequate experimental control or to the insensitivity of measurement.

The present report which describes two experiments and one control study is predicated on the hypothesis that both inadequate experimental con-

---

\* Received in the Editorial Office on July 3, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was carried out during the tenure of a Predoctoral Research Fellowship #MF 17,301 awarded to the first author. The authors thank Dr. Ellsworth Bourque of Rutland V.A. Hospital for his help in obtaining the Ss for this study.



trol and insensitivity of measurements can account for the earlier findings of no differences, from which discontinuity in learning was inferred.

## B. EXPERIMENT 1

The first experiment attempts to introduce more careful experimental controls for differential item difficulty, hypothesized to be absent in the earlier studies. Items from a pool of items, that are missed after three learning trials, may well be more difficult than items randomly selected from the total pool of items. Thus, in the Rock and Heimer study (7), the new items, i.e., those not previously presented, were selected from the total pool. These new items were then compared with items missed two times. It is reasonable to assume that items from a pool of items, missed two times are, on the average more difficult, than the other items in the pool. The greater difficulty of the missed items might be masking the facilitating effect of the three previous presentations. There are two ways of trying to control for item difficulty: (a) match the items by some independent assessment procedure, and test the continuity-discontinuity question by using a repeated measurement design, with each *S* as his own control;<sup>2</sup> and (b) match the *S*s for learning ability on some independent task and test each matched pair of *S*s for the same missed items, except one of the matched *S*s will have had three previous learning trials vs. none for the other matched *S*. In this experiment, the second control method, i.e., the matched group design, was utilized. While this procedure does not remove the possibility of idiosyncratic differences among *S*s in the difficulty of the various items for any *S*, it is based upon the reasonable assumption that those items which are most difficult for one member of a given population will most likely be difficult for another member of the same population.

Besides the control for item difficulty, the first experiment was also designed to investigate the influence of length of list on the continuity-discontinuity question. Reed and Riach (5) suggested that too short a list makes the learning task so easy for all *S*s that the continuity effect may be masked.

### 1. Method

*a. Subjects.* Twenty-eight male convalescent patients from a general medical V.A. hospital composed the sample (14 in the experimental group and 14 in the control group). There were nine additional *S*s who also participated in some portions of this study but who were not included in the two populations. The experimental design was such that all *S*s in the experimental

<sup>2</sup> After this study had been completed, Battig (1) reported an experiment in which he utilized this method and found evidence for the continuity view.

group who made less than 10 errors on the 60 item list (see materials below) were dropped from the experiment. Four of the nine *Ss* made less than 10 errors on this list. Four experimental *Ss* did not complete the learning task because they complained of some physical ailment. One *S*, randomly selected for the control group, could not complete the experiment because of a headache.

*b. Materials.* (1) Stimulus Presentation Cards. The paired-associates on each presentation card was a nonsense syllable (100 per cent and 93 per cent association value selected from Glaze's list) paired with a two-digit number. The nonsense syllables and numbers composed respectively the left and right hand members of the paired-associates. Each pair was typed in capital letters on a  $3 \times 5$  card. There were two sets of presentation cards, one of 20 items, one of 60 items. (2) Recognition Test Cards. Each recognition test ( $3 \times 5$ ) card had one of the nonsense syllables typed in the upper-middle portion of the card. Directly below the nonsense syllable, typed in two symmetrical columns, were four two-digit numbers, one of which was the "correct" right hand number. A Recognition Test Card was made for every Presentation Card. (3) Other Material. A black cardboard partition was used to block the *Ss'* view of the data sheet and the cards.

*c. Procedure.* All *Ss* were first given three presentations of the preliminary 20 item list, under instructions to learn as many of the pairings as possible. Each presentation card with a paired-associate was exposed manually by the *E* for two seconds, followed by a three-second rest interval. After a presentation of the complete list, a recognition test trial was given. Approximately 10 seconds was the interval between the end of the presentation card list and its test trial. During the recognition trials the *Ss* were allowed to examine the test cards until an answer was given. To randomize order, the presentation and test cards were shuffled after every complete presentation and test. After the third test trial, *S* was excused and asked to return the following week.

Each *S* assigned to the experimental group was matched with an *S* assigned to the control group, on the basis of the number of errors on the second and third test trials on the 20 item list. Assignment to each group was random within each pair of matched *Ss*.

In the second session, the experimental *S* was given three presentations of the 60 item paired-associate list, followed by two recognition test trials. After a short rest interval (about five minutes), the experimental *S* learned those items he had missed on *both* of these test trials to a criterion of one perfect trial. If an experimental *S* made less than 10 errors on the 60 item



list he was dropped from the experiment; if the *S* made more than 25 errors on the 60 item list, 25 items from the total error list were randomly selected to compose the final criterial learning list.

The second part of the experiment for the control *S* of the matched pairs was to have three presentations and two recognition test trials of those items correctly identified by his matched experimental *S* from the original 60 item list. These presentations partially controlled for practice, pro-active inhibition, or changes in motivation. However, the control *S*s, whose matched experimental *S*s made many errors (17-25) on the original 60 item list, had lists ranging from 35 to 43 items, while the control *S*s whose matched experimental *S*s made few errors (10-16) were presented with lists ranging from 44 to 50 items. The third part of the experiment for the control *S* was learning those items missed by his matched experimental *S* on the original 60 item list to the criterion of one perfect trial.

The most important procedure is in this third step. The experimental *S* is learning a list of items he has seen three times and missed, while the control *S* is learning the *same* list as his matched experimental *S*, but the control has *not* had these items presented to him previously. If learning is a continuous incremental process, then the experimental *S*s, having had previous experience with the items, should show faster learning than the control *S*s.

Three measures of learning were employed: trials to criterion, number of errors on the first criterial test trial, and the total number of errors in reaching criterion.

## 2. Results

*a. Number of errors on first critical test trial.* The error data were converted by a square root transformation to achieve homogeneity of variance. The results of the analysis of variance of the number of errors on the first criterial test trial reveals: (a) a significant interaction ( $F = 13.65$ ;  $df = 1, 12$ ;  $p < .01$ ) between the two conditions List Length and Experimental vs. Control; (b) a significant difference ( $F = 5.48$ ;  $df = 1, 12$ ;  $p < .05$ ) between the experimental and control groups; and (c) a significant difference ( $F = 63.31$ ;  $df = 1, 12$ ;  $p < .01$ ) between long and short lists. The long-short list dichotomy was based on a rank ordering of the length of the final list learned to criterion by the experimental group of *S*s. The seven longest lists which ranged from 17 to 25 errors were included in the "Long List" category, and the seven shortest which ranged from 10-16 errors in the "Short List" category. Under the short list condition, the difference between the mean number of errors by the experimental group (1.86) and by the

control group (5.57) was significant ( $p < .01$ ). However, under the long list condition the difference in mean errors between the experimental (10.57) and control (9.29) groups was not significant ( $p > .05$ ).

*b. Total number of errors to criterion.* Total number of errors by the experimental and control Ss in reaching criterion showed the same results as the analysis of the number of errors on the first criterial test trial. A significant interaction ( $F = 12.13$ ;  $df = 1, 12$ ;  $p < .01$ ) between list length and the experimental-control conditions was found. Under the short list condition, the difference between the mean number of errors to criterion of the experimental (3.86) and the control (9.29) groups was significant ( $p < .01$ ). However, under the long list condition, the difference in mean total errors of the experimental (29.00) and control (26.29) groups was not significant ( $p > .05$ ).

*c. Trials to criterion.* When trials to criterion were compared, no significant differences ( $F = 1.63$ ;  $df = 1, 12$ ) between the experimental and control groups were found for either the long list or short list conditions.

*d. Analysis of item difficulty.* To determine whether differential difficulty among items could be demonstrated, as posited in the introduction, a chi square was computed. The distribution of errors made on the first two test trials following the three presentations of the 60 item list for all 22 Ss who had completed the first part of the procedure for the experimental group was analyzed. A chi square of 84.69 ( $df = 59$ ;  $p < .025$ ) was found.

### 3. Discussion

The analysis of the number of errors on the first criterial test trial showed that the experimental group, consistent with the continuity view, made significantly fewer errors than did the control group. Thus, when a control for differential item difficulty is introduced, the continuity effect is evident. Also, the two error measures (number of errors on first criterial test trial and total errors made in reaching criterion) showed a significant interaction between list length and the experimental-control condition. With both measures, under the short list condition the experimental group made significantly fewer errors than did the control group, but under the long list condition no significant difference between the experimental and control groups was found. Originally, it had been expected that the opposite relationship would be evident, that is, the continuity effect to be more apparent under the long list condition than under the short list. It is possible that these results may be attributed to a differential fatigue or motivation effect. For the short list Ss, there was very little difference in the total amount of time re-



quired by the experimental and control Ss to complete the experiment—about 60 minutes. However, there was a significant time difference between experimental and control Ss under the long list condition. The average amount of time for a long list experimental S to complete the second session was from 90 to 100 minutes. The average time for a long list control S was approximately 70 minutes. The time difference can be attributed to the three presentations of the second testing session described in the procedures. The experimental Ss were presented with a 60 item list vs. only a 35 to 43 item list for the control Ss. Since it can be assumed that fatigue increased with the amount of time spent on the total task, it seems reasonable to conclude that the long list experimental group had more fatigue than did the long list control group. The apparent absence of evidence of continuity on the long list condition may thus be due to this postulated detrimental effect of fatigue, which masked the posited, positive effect of the three previous presentations.

The significant chi square found for the distribution of errors made by the experimental Ss on the 60 item list indicates that some items were more difficult to learn than others. This finding increases the tenability of the hypothesis that item difficulty was masking the results in the Rock; and the Rock and Heimer studies, as well as the studies by Estes and his associates.

### C. EXPERIMENT 2

While the first experiment dealt with control questions in the paired-associate learning paradigm, the second experiment attempts to utilize a more sensitive measure of learning, that is, latencies. Again, if learning is an incremental process, as the continuity hypothesis maintains, then items that have been presented several times but missed should be responded to more quickly than items that have not been presented previously.

#### 1. Method

*a. Subjects.* Twelve of the 28 Ss used in the first study composed the sample for the second experiment.

*b. Material.* Forty-seven paired-associate learning items, consisting of 87 per cent and 93 per cent association value nonsense syllables (from Glaze's list) and three digit numbers, were the stimulus material for the study. Thirty-two pairs made up the "previously presented" list, and 15 items the "not previously presented" list. There were two different types of cards used in the experiment:

(1). *Training cards.* Each syllable-number paired-associate was typed

on a separate  $3 \times 5$  card in capital letters. The nonsense syllables and numbers were the left and right hand members of the paired-associates respectively. Training cards were made only for the 32 items in the "previously presented" list.

(2). *Recognition cards.* Each recognition card contained a three-letter nonsense syllable in the upper-middle section of the ( $3 \times 5$ ) card and three three-digit number choices typed in a single column directly below the nonsense syllable. A recognition test card was made for each of the items in the "previously presented" list. Also, 15 other test cards, similar to the above, composed the "not previously presented" list. For these 15 Recognition Test Cards there were no training cards.

The Ss' response latencies to the test cards were measured in 1/10 seconds with a stop-watch.

c. *Procedure.* The Ss were given two presentations of the 32 item "previously presented" list. The inter-item and inter-presentation time intervals employed in Experiment 1 were used in this experiment. Following the second presentation, two recognition test trials were given. To randomize order, E shuffled the cards following every complete presentation and test trial. The test cards of those items missed on both test trials were removed from the deck. For every missed item ("previously presented") a new item ("not previously presented"), which S had never seen, was introduced. The "not previously presented" and "previously presented" test cards were then shown to the Ss for another test trial. The E recorded the Ss' answers and their speed of responding to each card. If S inquired whether he had seen the cards with the new items before, the E told him that he had seen all the cards earlier in the experiment.

## 2. Results

A logarithmic transformation of the latencies was computed to achieve homogeneity of variance. The results of the analysis of variance showed that the mean log latency (.712) for the "previously presented" items was significantly shorter ( $F = 45.75$ ;  $df = 1, 11$ ;  $p < .01$ ) than the mean log latency (.945) for the "not previously seen" items.

## 3. Discussion

The results of the second experiment also support the continuity hypothesis. The average latency for the "not previously presented" cards was significantly longer than the mean latency of the "previously presented" items. Furthermore, for all 12 Ss the latencies were shorter for the "previously pre-



sented" item choices; there were no reversals of this trend. If latency is a measure of learning, the greater speed with which the *Ss* responded to the "previously presented" items implies that the *Ss* did learn something about the paired associations from the previous two presentations. While other traditional measures of learning may fail to tap this relatively weak learning, latencies appear to be a more sensitive indicator.

Although the differences in speed of responding were significant, the use of latencies may involve an assumption not warranted, i.e., *S* has learned a correct response. Thus, *S* may not have learned the correct association but still responds quickly to the cards because of the presence of an incorrect association. If this hypothesis of incorrect association is valid, it would be expected that the *Ss* do less well than chance on the "previously presented" items during the third test trial. Since *Ss*' performance on the "not previously presented" cards can be considered a measure of chance guessing, a comparison of *Ss*' performance on the "previously presented" and "not previously presented" items supplies an empirical test of the "incorrect association" hypothesis. The difference (.25) between the means of the number of correct, previously seen items (1.75) and not previously seen items (2.00) on the third test trial is far short of statistical significance ( $F = .27$ ;  $df = 1, 11$ ). Thus, the number of correct responses on the third test trial shows no evidence for an "incorrect association" factor.

#### D. EXPERIMENT 2 A

This experiment was to provide a control for the second study. In employing latencies there is a question of familiarity with, or recognition of, the first member of the pair of the paired associates. One might reason that the faster latencies to the previously seen items were a function of the *Ss*' familiarity with the nonsense syllable that appeared at the top of every test card rather than with its pairing. That is, the *Ss* took less time to respond to the previously seen cards, because they "recognized" the nonsense syllables, rather than because they had learned the association of the previously seen card pairs. The third experiment was designed to investigate this possibility of familiarity effecting the speed of responding in the test series.

##### 1. Method

*a. Subjects.* Fourteen undergraduate students served as volunteer *Ss* for the third experiment.

*b. Materials.* Twenty 87 per cent and 93 per cent association value nonsense syllables and two-digit numbers were used in constructing the three

types of cards used in this experiment. These 20 syllables were selected randomly from the 47 in Experiment 2.

(1). *Presentation cards.* Ten of the 20 nonsense syllables were utilized in constructing the presentation series. Each of the 10 syllables was typed in capital letters on a separate  $3 \times 5$  card.

(2). *Training cards.* Each of the 20 nonsense syllables along with its correct two-digit number association were typed in capital letters on separate  $3 \times 5$  cards. The nonsense syllables and numbers made up respectively the left and right hand member of the paired-associate.

(3). *Test cards.* A separate test card was constructed for each of the 20 nonsense syllables. The syllables were typed in capital letters in the upper-middle part of the cards. Directly below the syllables four number choices, including the number paired with the syllable on the training cards, were typed in two symmetrical columns.

As in the previous experiments a cardboard partition was used to block the Ss' view of the score sheet. The latencies were measured in 1/10 seconds by a stop-watch.

c. *Procedure.* The Ss were told that a series of cards was to be shown to them. Their task was to read and familiarize themselves with the nonsense words on the cards. Three exposures of the 10 presentation cards was then given. The cards were shuffled after the first and second exposures of the whole list. The inter-item and inter-presentation time intervals employed in the first two experiments were also utilized in this experiment.

Upon completion of the third presentation, the directions that were used in the first experiment were read to the Ss. When S understood the nature of the experimental task, he was given one presentation of the training list. Following the training list, the test cards were presented once to the Ss. S's answer and his speed of responding were noted for every card.

## 2. Results

A logarithmic transformation of the latency scores was computed to approximate a normal distribution. An analysis of variance of the log of the latencies was then computed. The results revealed no significant difference ( $F = 0.076$ ,  $df = 1, 13$ ) between the log latencies of the familiar items (the 10 syllables exposed three times before the learning trial) and the nonfamiliar items (the 10 syllables not exposed before the learning trial). The mean log latency of the familiar items (0.7939) was slightly longer than the mean log latency of the non-familiar items (0.7863). Seven of the



14 Ss showed shorter log latencies to the familiar items and the other seven Ss shorter latencies to the non-familiar items.

### 3. Discussion

The results of the third experiment revealed no evidence for familiarity as a factor in the latencies to the test cards in the second experiment. If familiarity was operative, the latencies to the familiar items should have been significantly shorter than the latencies to the non-familiar items. This difference as a function of familiarity was not found.

### E. SUMMARY AND CONCLUSIONS

This study was a theoretical and empirical investigation of the recent evidence on the continuity-discontinuity views of learning. In particular, this study raised questions about the findings and conclusions of Rock; Rock and Heimer; Estes; and Estes, Hopkins and Crothers. The findings in the present investigations support the continuity point of view. Evidence was produced which indicated that the data from which discontinuity had been inferred by the previous investigators could be attributed to: (a) methodological factors; (b) differential item difficulties; and (c) insensitivity of their measurements of learning.

In the first of three experiments a theoretical and empirical analysis concerning differential item difficulty was attempted. With a control for item difficulty included in the experimental design, the results of the experiment revealed evidence for the continuity hypothesis. In addition, a highly significant interaction between list length and the experimental control factors was found. When the Ss had to learn short paired-associate lists, there was evidence for continuity. When the Ss had to learn a long list, the apparent facilitating effect of repetitions disappeared. Differences in the effects of the two list lengths was explained by positing an interaction of fatigue or decreased motivation which had a detrimental effect on the experimental Ss in the long list group.

In the second experiment, latencies were employed as a more sensitive indicator of weak associations for incorrectly associated items. It was hypothesized that if the continuity view is tenable Ss should respond more quickly to items they have seen although incorrectly associated than to items they have never seen. The results confirmed this expectation. The Ss responded more quickly to the previously exposed items. The criticism of using latencies as indicators of correct weak associations, namely the existence of wrong associations was discussed and explored. That is, the Ss may respond more quick-

ly to the repeated cards not because of a weak association but because of a well-learned wrong association. No such evidence for this wrong association hypothesis was found.

The third experiment attempted to investigate the possibility of artifacts underlying the evidence for continuity; that is, the low latencies to the previously presented items was a function of the recognition of or the familiarity with the stimulus member of the paired-associate item rather than to the association of the stimulus and response number of the pair. The results of the third study produced no evidence that more familiar syllables are responded to more quickly, and the conclusions drawn from the data of the second experiment as support of the continuity view is tenable.

#### REFERENCES

1. BATTIG, W. F. Paired-associate learning under simultaneous repetition and non-repetition conditions. Paper presented at EPA, Atlantic City, N.J., April, 1962.
2. CLARK, L. L., LANSFORD, T. G., & DALLENBACH, K. M. Repetition and associative learning. *Amer. J. Psychol.*, 1960, **73**, 22-40.
3. ESTES, W. K. Learning theory and the new mental chemistry. *Psychol. Rev.*, 1960, **67**, 207-223.
4. ESTES, W. K., HOPKINS, B. L., & CROTHERS, E. J. All-or-none and conservation effects in the learning and retention of paired associates. *J. Exper. Psychol.*, 1960, **60**, 329-339.
5. REED, J. C., & RIACH, W. D. The role of repetition and set in paired-associative learning. *Amer. J. Psychol.*, 1960, **73**, 608-611.
6. ROCK, I. The role of repetition in associative learning. *Amer. J. Psychol.*, 1957, **70**, 186-193.
7. ROCK, I., & HEIMER, W. Further evidence of one-trial associative learning. *Amer. J. Psychol.*, 1959, **72**, 1-16.

*Department of Psychology*  
*Clark University*  
*Worcester 10, Massachusetts*



## CONSISTENCY BETWEEN ATTITUDINAL AFFECT AND SPONTANEOUS COGNITIONS\*<sup>1</sup>

*Ohio State University and Yale University*

---

MILTON J. ROSENBERG AND PHILIP K. OLTMAN

---

### A. INTRODUCTION

A "structural approach" to social attitudes has been described, and supporting data reported, in previous publications (5, 6, 7, 8).

At the core of this approach are two main propositions. The first is that stable social attitudes have an internal structure in which their affective and cognitive components stand in consistent relationship to one another. Typical of evidence supporting this proposition are some findings on the relationships between the affective and cognitive aspects of the attitude toward desegregation. For example it was found (5) that persons strongly in favor of housing desegregation held beliefs to the effect that desegregation clearly facilitates the attainment of certain of their positive values (e.g., "all human beings having equal rights"; "people being well educated"; etc.) and prevents the actualization of certain negatively valued conditions (e.g., "people sticking to their own groups"; "achieving superiority over others"). Conversely, persons strongly opposed to desegregation were found to believe that it fosters the attainment of negatively valued conditions and prevents the attainment of positively valued conditions. In the case of more moderate positive or negative attitudinal affects the associated beliefs were found to relate the attitude object to less important values or, if to important values, then with less confidence as to the existence of clear-cut instrumental relationships between the attitude object and the values in question.

A second major proposition is a more dynamic one: it asserts that the disruption of the internal consistency of attitudes through the alteration of either the affective or cognitive component will, if the resulting tension is of sufficient intensity, be followed by the re-establishment of consistency through attitude reorganization. Some studies have been reported in which alteration of the cognitive component has led to change in attitudinal affect

---

\* Received in the Editorial Office on July 3, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was carried out under Contract 609(27) with the Group Psychology Branch of the Office of Naval Research while the senior author was a member of the Psychology Department at Yale University.

(2, 4) and other studies (6, 8) have shown that alteration of the affective component leads to change in the cognitive component.

In most of the aforementioned studies the affective aspect of attitudes was measured with a 16-point scale running from extreme negative evaluation of the attitude object to extreme positive evaluation. In all of these studies the cognitive aspect of attitudes was measured with a device requiring *S* to judge some 15 to 30 "value terms" (e.g., "all human beings having equal rights"; "having interesting work to do"; "being allowed to maintain the privacy of one's opinions and beliefs"; etc.). Each value term was judged both for its "importance" as a positive or negative goal and for the degree to which its attainment was believed to be blocked or fostered by the attitude object in question.<sup>2</sup>

For the value importance measure the *S* assigned each value item to a position on a 21-category scale ranging from "gives me maximum satisfaction" (+10) through "gives me neither satisfaction nor dissatisfaction" (0) to "gives me maximum dissatisfaction" (-10). For the second judgment *S* rated each value on a 11-category scale. Category +5 represented "complete attainment" of the value through the instrumental agency of the attitude object. (Thus if, in the study concerned with housing desegregation, *S* believed that the institution of this policy would lead to the complete attainment of "all human beings having equal rights" he would assign that value item a scale position of +5.) Category 0 represented "neither attainment nor blocking" and category -5 represented "complete blocking" of the value through the effects of the attitude object. The intermediate categories on either side of 0 represented intermediate degrees of positive and negative instrumentality.

The main index of the cognitive component of the attitude was obtained by algebraically summing the importance-instrumentality products for each of the values. Thus when a positive value (rated, say, as +7) was perceived to be attained through the agency of some attitude object such as housing desegregation (to the extent, say, of a rating of +3 on "instrumentality"), the product was +21. If the same value were rated as "blocked" (say a rating of -3 on "instrumentality") the resultant product would be -21. Similarly, a *negative* value would yield a negative product if multiplied

<sup>2</sup> The attitude objects employed were proposed social actions or policy changes. For other types of objects other types of cognitions would be more relevant; i.e., cognitions about the objects "attributes" and their positive or negative value, or cognitions concerned with relationships of identity and similarity rather than with instrumentality.



by a rating of positive instrumentality and a positive product if multiplied by a rating of negative instrumentality. The single algebraic quantity obtained by summing all such products was taken to represent the total import of the *S*'s pattern of beliefs about the attitude object's value-attaining and value-blocking powers.

In addition to its employment in all of the aforementioned studies this technique for the measurement of the cognitive aspect of attitudes has also been used by Axelrod (1), Miller (3) and Scott (9) and the data from their studies agree with those of Rosenberg (5) in confirming the proposition that the affective and cognitive aspects of attitudes are closely and consistently co-ordinated with one another.

However it can be argued that the method outlined here for the assessment of the cognitive content of an attitude is not completely satisfactory; that in asking *S* to characterize the relationship between the attitude object and certain standardized value terms we are eliciting essentially *new* judgments which are governed by the response set established through prior measurement of his attitudinal affect. While additional evidence (8) suggests that this is not usually the case, it would be useful to circumvent the problem if possible. It would also be desirable to adapt the technique for the measurement of attitudinal cognition to the requirements of the kind of interview approach usually employed in "public opinion" research. Both these purposes could be served by allowing the values that are phenomenally salient to the individual's attitude to emerge spontaneously. Essentially, this is what is done in the "open-ended" interview method of attitude assessment; but that method does not usually permit the kind of quantification achieved with the technique used in the aforementioned studies.

## B. METHOD

In an attempt to combine the best features of both approaches we developed a new procedure involving the following sequence of steps:

1. *S*'s attitudinal affect toward a given object (in the case of the present study, a policy change) is elicited by use of a sixteen point scale, running from "extreme pro" to "extreme anti."
2. He is then asked to give his thoughts on the policy change, stressing particularly his beliefs about its likely "consequences." Typical recorded statements are: "it (the policy change) will make it harder to keep up with your studies"; "it will help in making new friends."
3. The interviewer transcribes verbatim all the positive and negative value-like terms (consequences, goals, ends, etc., as for example, "keeping up

with studies" and "making new friends," above) spontaneously mentioned by *S*.

4. When *S* has finished speaking the interviewer asks him to rate each of his own, spontaneously produced, value-like terms (on a scale from  $-3$  to  $+3$ ) for the degree to which it is a "desirable" or "undesirable" end, and also (on a scale from  $+1$  to  $+3$ ) for the degree to which it would probably be "fostered" through the instrumental effects of the attitude object.

5. For each spontaneously elicited value term the two ratings are multiplied and the sum of the resultant products is computed. In effect we have obtained from *S* the same type of attitudinal cognition score used in earlier studies (2, 5, 6, 8) but now based only upon the value terms which for him are salient to his attitude and which are spontaneously associated with it.

### C. RESULTS

Using this procedure with forty-four Yale undergraduates on each of four different attitude issues, it was possible to retest the basic proposition to the effect that the affective and cognitive aspects of attitudes are consistent with one another. Successful and strong replication of the earlier findings was obtained. The accompanying table presents the results for each of the four attitude issues respectively. In each analysis the distribution of affect scores (degree to which *S* is "in favor" or "opposed" to an advocated change in some aspect of university life) is broken into approximately equal quarters and the cognition scores into approximately equal thirds, thus duplicating the form of analysis reported by Rosenberg (5).

As Table 1 shows, in each of the four attitude areas the *Ss'* affective responses are significantly and consistently related to their cognitive responses. For three of the attitudes investigated the relationship between the indices of affect and cognition are significant at a probability of less than .001. In the fourth case the obtained probability is less than .01. Extreme positive and negative affect scores are associated with extreme positive and negative scores respectively on the index of attitudinal cognition; and more moderate affect scores are associated with more moderate scores on the cognitive index. Examination of the individual cases reveals: (a) that this pattern of covariance is due to the fact that cognitions associated with extreme positive or extreme negative affective responses involve values of greater intensity, and instrumental relationships of greater certainty, than those associated with more moderate positive or negative affects; (b) that cognitions associated with positive affects are in the main ones which relate the attitude object to the attainment of positive values while in the case of negative affects the



TABLE 1  
THE RELATION OF ATTITUDINAL AFFECT AND COGNITION ON FOUR UNIVERSITY ISSUES

A. How would you feel toward the idea of <i>having Yale University become coeducational?</i>			
Affect scores	Cognition scores		
	+20 to +5	+4 to -4	-6 to -29
+8 to +4	10	4	0
+3 to -1	3	6	2
-2 to -5	2	3	6
-6 to -8	0	2	6
$X^2 = 22.7; p < .001$			
B. How do you feel toward the idea of <i>discontinuing the secret societies at Yale?</i>			
Affect scores	Cognition scores		
	+19 to 0	-1 to -5	-6 to -21
+7 to +1	10	1	1
-1	3	4	1
-2 to -4	5	3	4
-5 to -8	1	3	8
$X^2 = 13.5; p < .01$			
C. How do you feel toward the idea of <i>de-emphasizing inter-collegiate football at Yale?</i>			
Affect scores	Cognition scores		
	+12 to -4	-5 to -10	-11 to -30
+4 to -3	7	0	1
-4 to -5	5	6	2
-6 to -7	1	7	2
-8	1	2	10
$X^2 = 30.5; p < .001$			
D. How would you feel toward <i>expanding the size of Yale?</i>			
Affect scores	Cognition scores		
	+27 to +2	+1 to -3	-4 to -36
+8 to +3	8	3	0
+2 to -2	3	8	1
-3 to -5	3	3	6
-6 to -8	1	1	7
$X^2 = 25.1; p < .001$			

associated cognitions relate the attitude object to the attainment of negative values.

#### D. SUMMARY AND CONCLUSIONS

A new method for eliciting and quantifying the cognitive aspect of social attitudes was described. Data were reported concerning the relationship between an index of attitudinal cognition based upon this procedure and a separate index of attitudinal affect. These data seem relevant to the field of attitude study in two ways: they provide useful additional confirmation of the proposition that the affective and cognitive components of attitudes are typically organized in close consistency with one another; they help validate

a new method of attitude measurement that may have wide utility both in experimental studies and in sample surveys of social attitudes.

## REFERENCES

1. AXELROD, J. The relationship of mood and mood shift in attitude. Technical report to the Office of Naval Research. Univ. Rochester, 1959.
2. CARLSON, E. R. Attitude change and attitude structure. *J. Abnorm. & Soc. Psychol.*, 1956, **52**, 256-261.
3. MILLER, L. Subjective probability estimates as a function of mood and attitude structure. Unpublished doctoral dissertation. Univ. Rochester, 1960.
4. PEAK, H. The effects of aroused motivation on attitudes. Technical report to the Office of Naval Research. Univ. Michigan, 1959.
5. ROSENBERG, M. J. Cognitive structure and attitudinal affect. *J. Abnorm. & Soc. Psychol.*, 1956, **53**, 367-372.
6. ———. Cognitive reorganization in response to the hypnotic reversal of attitudinal affect. *J. Pers.*, 1960, **28**, 39-63.
7. ———. A structural theory of attitude dynamics. *Pub. Opin. Quart.*, 1960, **24**, 319-340.
8. ———. An analysis of affective-cognitive consistency, 1960. In M. J. Rosenberg, *et al.*, *Attitude Organization and Change*. New Haven: Yale Univ. Press, 1960.
9. SCOTT, W. A. Cognitive consistency, response reinforcement, and attitude change. *Sociometry*, 1959, **22**, 219-229.

Department of Psychology  
The Ohio State University  
1945 North High Street  
Columbus 10, Ohio



## STIMULUS SIZE AND RORSCHACH RESPONSES\*

*University of Utah, College of Medicine*

---

IJA N. KORNER, ROGER B. ALLISON, JR., AND MAX D. ZWANZIGER

---

### A. INTRODUCTION

In general, variations in stimulus properties of the Rorschach cards have not modified the nature of resultant percepts. Baughman's review (3) suggested that the majority of Rorschach investigators have been concerned with the role color played in Rorschach responses and have found that color had little influence upon percepts. Balloch (1) varied the shading contrast of ink blots and failed to find significant variations in physiological or verbal responses to the cards. Baughman (2) conducted a similar but more comprehensive study of the blots in which modifications in color, shading, figure-ground and form characteristics were introduced. He found that the percepts were not particularly changed although the stimulus characteristics were altered considerably.

It thus appears that Rorschach responses are relatively invariant under marked variations in stimulus properties. One characteristic of interest is the physical size of the Rorschach stimulus, which has not been studied directly. Group techniques (4, 5, 6, 7, 8) in which slides of Rorschach cards are projected on a screen are not impressive as a valid procedure for investigating the effect size may have upon percepts. In a culture which abounds with motion picture experiences, as well as familiarity with the traditional  $7 \times 9 \frac{1}{2}$  inch versions of ink blots, perceived physical size could well be constant regardless of the size of the projected material. The present paper is concerned with the influence that physical size has upon Rorschach responses.

### B. METHOD

The stimulus material for the study consisted of two sets of Rorschach Cards I, III, and VI. One set was the standard clinical size of  $7 \times 9 \frac{1}{2}$  inches and the second was  $6 \times 8$  foot photographic enlargements. The latter set required mounting on three walls of an adequately sized room.

Ss were 38 university students, 18 to 24 years of age, who were enrolled in an introductory psychology course and were required to take part in the

---

\* Received in the Editorial Office on July 18, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

experiment. Both sets of Rorschach material were individually administered by standard clinical procedures to each *S*. When viewing the enlarged blots, *Ss* were seated six feet from the cards. Head movements were minimized through the use of a chin rest which was adjusted to each *S*. The order of presentation of the small and king size sets was counterbalanced to minimize order effects. A verbatim account of responses and the inquiry was kept.

In addition to scoring by the Beck system, the data were scored for: (a) the total number of words, which included comments and conversational remarks as well as responses; (b) the number of percepts judged larger than a two foot cube; (c) the number of percepts judged smaller than a two foot cube; and (d) the number of threatening responses.

### C. RESULTS AND DISCUSSION

The total number of responses evoked by the small cards was 402 as compared with 439 responses for the king size cards. The difference in the number of responses was not statistically significant by Wilcoxon's Test for Paired Replicates, which was the statistical test used throughout the analysis. The following scoring categories also failed to support a size difference: Total Number of W, D, Dd, F+, F—, Color (C, CF, FC), M, Human (H, Hd), Animal (A, Ad), and Popular responses, and also Reaction Time. Total time, however, was significantly greater at the .01 level of confidence for the king size cards: 310 minutes were taken in responding to the king size cards as compared to 239 minutes for the small cards. Thus, the traditional scoring categories remained relatively invariant to changes in the physical size of the Rorschach stimuli. The only exception to this conclusion was that *Ss* took significantly more time responding to the king size cards than to the small cards.

The total number of words used during the test period and inquiry was significantly greater for the king size cards (5594 words) than the small size cards (4376); the level of confidence was again .01. Thus *Ss* not only took significantly more time responding to the king size cards but used more words in so doing. Because of the dependency suspected between these two variables a product moment correlation coefficient was computed. The obtained correlation of .76 implies that the two variables which were significantly influenced by changes in stimulus size—word count and total time—may indicate the possibility of a more basic parameter. As a hypothesis for further research, we suggest that the underlying parameter is a function of the difficulty involved in the perceptual organization of the larger blots and the anxiety or discomfort aroused during the period of gaining closure.



The number of percepts judged larger or smaller than a two foot cube for the two stimulus sizes did not yield differences significant at the .01 level. However, the difference in the number of "small" responses to the king size cards (291), compared with the small cards (242), did reach the .07 level of confidence and may warrant further study. A tendency to use "smaller" rather than "larger" conceptualizations for the king size cards might represent a true difference. The king size cards could overwhelm the *S* and produce a feeling of smallness which was then projected into the response.

The number of threatening responses (i.e., "a mean face looking right at you," "a small jungle animal jumping at something from a tree") was essentially the same for the two sets of stimuli: 34 for the small cards and 33 for the king size cards. If more anxiety was aroused by the king size cards, as postulated above, it did not reach a sufficient level to be demonstrated in the responses.

#### D. SUMMARY

In order to explore the influence which physical size of Rorschach stimuli has upon responses, two sets of Rorschach blots were administered to 38 university students. The first set was the standard clinical size of  $7 \times 9 \frac{1}{2}$  inches. A king size edition of the same material, measuring  $6 \times 8$  feet, constituted the second set. In terms of the present study, it was concluded that variations in the physical size of the stimuli influence the total time and number of words used in responding. The nature of percepts represented by scoring categories of Beck's system, with the exception of total time, appeared to remain invariant to changes in stimulus sizes. One wonders if a general recommendation might be made to Rorschach researchers, who investigate the effects of color, shading, etc., to develop response categories outside the traditional measures if differences are sought. The Rorschach may then become more sensitive to the measurement of personality changes.

#### REFERENCES

1. BALLOCH, J. C. The effect of degree of shading contrast in ink blots on verbal response. *J. Exp. Psychol.*, 1952, **43**, 120-124.
2. BAUGHMAN, E. E. A comparative analysis of Rorschach form with altered stimulus characteristics. *J. Proj. Tech.*, 1954, **18**, 151-164.
3. ———. The role of the stimulus in Rorschach responses. *Psychol. Bull.*, 1958, **55**, 121-147.
4. HARROWER-ERICKSON, M. H., & STEINER, M. E. Modification of the Rorschach method for use as a group test. *Rorschach Res. Exch.*, 1941, **5**, 130-144.
5. HERTZMAN, M. A comparison of the individual and group Rorschach test. *Rorschach Res. Exch.*, 1942, **6**, 89-108.

6. LAZARUS, R. S. An experimental analysis of the influence of color on the protocol of the Rorschach test. *J. Personal.*, 1948, 17, 182-185.
7. ———. The influence of color on the protocol of the Rorschach test. *J. Abn. & Soc. Psychol.*, 1949, 44, 506-515.
8. LINDNER, R. M., & CHAPMAN, K. W. An eclectic group method. *Rorschach Res. Exch.*, 1942, 6, 139-146.

*Department of Psychiatry*

*University of Utah*

*Salt Lake City, Utah*